

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
23 May 2002 (23.05.2002)

PCT

(10) International Publication Number
WO 02/40679 A2

(51) International Patent Classification⁷: **C12N 15/31**,
15/62, 15/77, 1/21, C12P 13/04

(21) International Application Number: PCT/US01/43096

(22) International Filing Date:
15 November 2001 (15.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/248,219 15 November 2000 (15.11.2000) US

(71) Applicants and

(72) Inventors: **RAYAPATI, P. John** [US/US]; 601 North
Union Street, Monticello, IL 61856 (US). **CRAFTON,**
Corey, M. [US/US]; 1371 West Pershing Road, Decatur,
IL 62526 (US).

(74) Agents: **LUDWIG, Steven, R.** et al.; Sterne, Kessler,
Goldstein & Fox P.L.L.C., 1100 New York Avenue, N.W.,
Suite 600, Washington, DC 20005-3934 (US).

(81) Designated States (*national*): AF, AG, AI., AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU,
ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: NUCLEOTIDE SEQUENCES FOR TRANSCRIPTIONAL REGULATION IN CORYNEBACTERIUM GLUTAM-
ICUM

(57) Abstract: The invention relates to isolated polynucleotides from *Corynebacterium glutamicum* which are useful in the regu-
lation of gene expression. In particular, the invention relates to isolated polynucleotides comprising *C. glutamicum* promoters which
may be used to regulate, i.e., either increase or decrease, gene expression. In certain embodiments, isolated promoter sequences
of the present invention regulate gene expression through the use of exogenous or endogenous induction. The invention further
provides recombinant vectors and recombinant cells comprising isolated polynucleotides of the present invention, preferably in op-
erably associated with heterologous genes. Also provided are methods of regulating bacterial gene expression comprising growth
of a recombinant cell of the present invention. In particular, the present invention provides methods to regulate genes involved in
amino acid production comprising growth of a recombinant cell of the present invention. In certain embodiments, the present in-
vention provides methods of regulating gene expression in bacteria, particularly *Corynebacterium* species, especially of the genus
Corynebacterium, comprising fermentation growth of a recombinant cell of the present invention, where metabolic concentrations,
temperature, or oxygen levels are manipulated to regulate gene expression.



WO 02/40679 A2

Nucleotide Sequences for Transcriptional Regulation in *Corynebacterium glutamicum*

Background of the Invention

Field of the Invention

5 The invention relates to the areas of microbial genetics and recombinant DNA technology. The invention provides DNA sequences, vectors, microorganisms, and methods useful for inducing and regulating the expression of genes, including those that are involved in amino acid biosynthesis, in bacterial cells.

10 *Background of the Invention*

 Coryneform bacteria are Gram-positive bacteria frequently used for industrial-scale production of amino acids, purines, and proteins. Although coryneform bacteria, particular *Corynebacterium* species, have been widely used for industrial purposes for many years, the techniques of molecular biology have
15 only recently been employed to augment the usefulness of these organisms in the production of amino acids and other products.

 One way to improve the productivity of a microbial strain is to increase the expression of genes that control the production of a metabolite. Increasing expression of a gene can increase the activity of an enzyme that is encoded by that
20 gene. Increasing enzyme activity can increase the rate of synthesis of the metabolic products made by the pathway to which that enzyme belongs. In some instances, increasing the rate of production of a metabolite can unbalance other cellular processes and inhibit growth of a microbial culture. The modified culture will make more product per cell, but will not be able to generate enough cells per
25 volume to show an improvement over the parent strain in a fermentor.

 Transcription is the process by which an RNA molecule is synthesized from a DNA template and occurs by the interaction of a multisubunit enzyme complex, known as RNA polymerase, with a DNA molecule. The RNA that is

-2-

synthesized by this process ultimately directs the production of protein products within the cell. In general, the rate at which RNA is synthesized from DNA, *i.e.*, the transcription rate, directly influences the level of synthesis of the corresponding protein product.

5 Promoters are DNA sequence elements that regulate the rate at which genes are transcribed. Promoters can influence transcription in a variety of ways. For example, some promoters direct the transcription of their associated genes at a constant rate regardless of the internal external cellular conditions. Such promoters are known as constitutive promoters. In many cases, however, a
10 promoter will direct transcription of its associated gene only under very specific cellular conditions. For example, promoters that turn off gene expression during the growth phase of a microbial culture, but turn on gene expression after optimal growth has been achieved can be used to regulate genes that control production of a metabolite. The new strain will have the same growth pattern as the parent
15 but produce more product per cell. This kind of modification can also improve titer (g product/liter) and yield (g product/g glucose). Nucleotide sequences have been identified that can be used to increase or decrease gene expression in *Corynebacterium* species. These regulatable promoters can increase or decrease the rate at which a gene is transcribed depending on the internal and/or the
20 external cellular conditions. Frequently, the presence of a factor, known as an inducer, can stimulate the rate of transcription from a promoter. Inducers can interact directly with molecules that, themselves, physically interact with the promoter or with DNA sequences in the vicinity of the promoter. Alternatively, the action of an inducer in stimulating transcription from a promoter may be
25 indirect. Whereas inducers function to amplify the level of transcription from a promoter, there is a class of factors, known as suppressors, that reduce or inhibit transcription from a promoter. Like inducers, suppressors can exert their effects either directly or indirectly.

Besides regulation through inducers and suppressors, certain promoters
30 are regulated by temperature. For instance, a the level of transcription from a

-3-

promoter may be increased when cells harboring that promoter are grown at a temperature that is greater than the optimum or normal growth temperature for that cell type. Similarly, there are promoters that will enhance gene expression in cells grown at temperatures below the normal growth temperature.

5 Promoters are found naturally in wild-type cells where they regulate the expression of specific genes. Promoters, however, are also useful as tools of molecular biology in that they can be isolated from their normal cellular contexts and engineered to regulate the expression of virtually any gene.

10 The use of regulatory sequences from *Escherichia coli* to control the expression of reporter genes in *Corynebacterium* have been documented. The lacI^q repressor and the *tac* promoter/reporter genes from *E. coli* were on plasmids that replicate in *Corynebacterium*. See, e.g., Morinaga, Y. *et al.*, *J. Biotechnol.* 5:305-312 (1987). In addition, Ben-Samoun *et al.*, *FEMS Microbiology Letters* 174:125 (1999), which is incorporated herein by reference in its entirety, disclose
15 the use of the *E. coli araBAD* promoter and the *araC* activator on a plasmid which replicates in *Corynebacterium glutamicum* cells to stimulate the expression of the GFPuv reporter gene only when L-arabinose is present in the growth medium. The authors acknowledge, however, that the level of expression from the *araBAD* promoter in *C. glutamicum* is 6.5 fold lower than that which was
20 observed in *E. coli*, the native species for the promoter.

U.S. Patent Nos. 5,693,781 and 5,762,299, each of which are incorporated herein by reference in their entireties, disclose the isolation of promoter sequences from the coryneform bacteria, *Brevibacterium flavum*. Sequences described in these patents were isolated on the basis of their ability to direct
25 expression of a reporter gene in *B. flavum* at a level that was greater than the expression level observed in *B. flavum* with the synthetic *tac* promoter. Also disclosed in U.S. Patent Nos. 5,693,781 and 5,762,299, each of which are incorporated herein by reference in their entireties, are *B. flavum* promoters capable of expressing a reporter gene in *B. flavum* cells when grown in medium
30 containing: (a) ethanol but not glucose, and vice versa; (b) glucose but not

fructose; and (c) glucose but not casein hydrolysates/yeast extract/glucose, and vice versa. The novel promoter sequences disclosed in the present application are different from those described in U.S. Patent Nos. 5,693,781 and 5,762,299.

A limited number of *C. glutamicum* promoters have been described to date. For example, the *C. glutamicum aceA* promoter is disclosed in Wendisch *et al.*, *Arch Microbiol.* 168:262 (1997) and in U.S. Patent No. 5,700,661 (where it is termed the isocitrate lyase promoter), each of which are incorporated herein by reference in their entireties. In both of these references, the *aceA* promoter was linked to a reporter gene in transformed *C. glutamicum* cells, and produced an extracellular protein, not a product of metabolic engineering. Expression of the reporter gene was found to be greater in *C. glutamicum* transformants that were grown in the presence of acetate than it was for transformants grown in the presence of glucose (Wendisch *et al.*, *Arch Microbiol.* 168:262 (1997)) or sucrose (U.S. Patent No. 5,700,661).

Similarly, the *aceB* promoter from *C. glutamicum* is disclosed in Wendisch *et al.*, *Arch Microbiol.* 168:262 (1997) and in U.S. Patent No. 5,965,391, which is incorporated herein by reference in its entirety. Both of these references describe transcriptional fusions consisting of the *aceB* promoter region linked to a reporter gene in *C. glutamicum* transformed cells. Expression of the reporter gene was found to be greater in *C. glutamicum* transformants grown in acetate-containing medium than it was for transformants grown in glucose-containing medium (Wendisch *et al.*, *Arch Microbiol.* 168:262 (1997)) or other carbon sources (U.S. Patent No. 5,965,391).

Reinscheid *et al.*, *Microbiology* 145:503 (1999), which is incorporated herein by reference in its entirety, discloses a transcriptional fusion between the *C. glutamicum pta-ack* promoter and a reporter gene (chloramphenicol acetyltransferase). *C. glutamicum* cells harboring the transcriptional fusion demonstrated enhanced reporter gene expression when grown in acetate-containing medium as compared to transformed cells that were grown in glucose-containing medium.

In Pátek *et al.*, *Microbiology* 142:1297 (1996), which is incorporated herein by reference in its entirety, several DNA sequences from *C. glutamicum*, identified on the basis of their ability to promote the expression of a chloramphenicol resistance reporter gene in *C. glutamicum* cells, are disclosed and compared to one another in an attempt to define a consensus sequence for *C. glutamicum* promoters.

There is clearly a need for a broader assortment of well-defined *Corynebacterium* species promoters than has been heretofore described. Such promoters would be useful in the constitutive and/or regulated expression of genes in coryneform cells. For example, a collection of *C. glutamicum* promoters, regulated by inexpensive carbon sources, would facilitate the industrial-scale production of amino acids and purines in *C. glutamicum* cells by enhancing the expression of genes that encode components of the biosynthetic pathways for the desired amino acids or purines. Likewise, a versatile array of coryneform promoters would be useful for the industrial scale production of heterologous polypeptides in *C. glutamicum* cells by stimulating the enhanced expression of genes encoding such heterologous polypeptides.

Summary of the Invention

The present invention relates to isolated polynucleotides that function as transcriptional regulators, in particular, promoters, in *Corynebacterium* species host cells, preferably *Corynebacterium glutamicum*, *Brevibacterium flavum*, or *Brevibacterium lactofermentum* host cells, even more preferably *C. glutamicum* host cells. These promoters are useful for regulating and enhancing the production of a variety of products in such host cells. Examples of products, the production of which may be enhanced in *Corynebacterium* species host cells as a result of the present invention, are amino acids, such as lysine; purine nucleotides, such as inosinic acid; and heterologous polypeptides. Since *Corynebacterium* species are especially useful for the industrial-scale production

of amino acids, purines, and polypeptides, use of the promoters of the present invention may greatly improve the yields of these products from *Corynebacterium* species host cells.

5 In one embodiment of the present invention, an isolated polynucleotide is provided, comprising a first nucleic acid, the sequence of which is selected from the group consisting of SEQ ID NO:4 through 22.

Further embodiments of the invention include an isolated polynucleotide comprising a first nucleic acid, the sequence of which is at least 90% identical, and more preferably at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% 10 identical, to a sequence selected from the group consisting of SEQ ID NOs: 4 through 22.

Additional embodiments of the invention include an isolated polynucleotide comprising a first nucleic acid, the sequence of which comprises about 6 contiguous nucleotides, preferably about 10 contiguous nucleotides, even 15 more preferably about 12, 15, 20, 30, 40, 50, 100, 150, 200, or 500 contiguous nucleotides, of a sequence selected from the group consisting of SEQ ID NOs: 4 through 22, wherein the first nucleic acid is capable of regulating transcription of a second nucleic acid, preferably as part of a promoter.

In one aspect of this embodiment, the promoter comprises any one or 20 more of the following genetic elements: a minus 10 (" -10") sequence; a minus 35 (" -35") sequence; a transcription initiation site; an enhancer region; and an operator region. Preferably the genetic elements are specific for *Corynebacterium* species, more preferably are specific for *Corynebacterium glutamicum*, *Brevibacterium flavum*, and *Brevibacterium lactofermentum*, and even more 25 preferably are specific for *Corynebacterium glutamicum*.

Additional embodiments of the invention include an isolated polynucleotide comprising a first nucleic acid at least 10 nucleotides in length, preferably at least 12, 15, 30, 50 or 150 nucleotides in length, which hybridizes under stringent conditions to a reference nucleic acid, or the complement thereof, 30 wherein the sequence of said reference nucleic acid is selected from the group

consisting of SEQ ID NOs: 4 through 22. The meaning of the phrase "stringent conditions" as used herein is described *infra*.

In further embodiments of the present invention, an isolated polynucleotide is provided, comprising a first nucleic acid as described above; and a second nucleic acid. According to this embodiment, the first nucleic acid comprises a transcriptional regulatory region, preferably, a promoter.

In one aspect of this embodiment, said second nucleic acid encodes one or more polypeptides. Preferably, the physical location of the first nucleic acid relative to the second nucleic acid is such that, under the appropriate conditions, the first nucleic acid regulates transcription of the second nucleic acid, thereby facilitating production of the polypeptide.

In another aspect of this embodiment, the second nucleic acid encodes one or more components of a purine biosynthesis pathway. One example of a purine biosynthesis pathway included in this aspect of the invention is the enzymatic pathway that results in the synthesis of inosinic acid.

In another aspect of this embodiment, the second nucleic acid encodes one or more heterologous polypeptides. The heterologous polypeptide may be one that is from a *Corynebacterium* species or one that is from a non-*Corynebacterium* species.

In one particularly preferred aspect of this embodiment, an isolated polynucleotide is provided, comprising a first nucleic acid as described above; and a second nucleic acid, which encodes a component of an amino acid biosynthesis pathway. Examples of amino acid biosynthesis pathways included in this aspect of the invention are the enzymatic pathways that result in the synthesis of L-glycine, L-alanine, L-leucine, L-methionine, L-phenylalanine, L-tryptophan, L-lysine, L-glutamine, L-glutamic acid, L-serine, L-proline, L-valine, L-isoleucine, L-cysteine, L-tyrosine, L-histidine, L-arginine, L-asparagine, L-aspartic acid, and L-threonine.

In a preferred aspect of this embodiment, the second nucleic acid encodes one of the following components of the L-lysine biosynthesis pathway:

aspartokinase, aspartate beta-semialdehyde dehydrogenase, diaminopimelate dehydrogenase, diaminopimelate decarboxylase, dihydrodipicolinate synthetase, dihydrodipicolinate reductase, or pyruvate carboxylase.

Another embodiment of the invention provides an isolated polynucleotide comprising a first nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 2, and a second nucleic acid which encodes a polypeptide which functions as a component of the lysine biosynthesis pathway. In this embodiment, the first nucleic acid regulates transcription of the second nucleic acid.

Yet another embodiment of the invention provides an isolated polynucleotide comprising a first nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 3; and a second nucleic acid which encodes polypeptide which functions as a component of an amino acid biosynthesis pathway, preferably a lysine biosynthesis pathway. In this embodiment, the first nucleic acid regulates transcription of the second nucleic acid.

Further embodiments of the invention provide isolated *Corynebacterium* chromosomes with a first nucleic acid integrated into the chromosome, the sequence of which is one of the following: a sequence which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:1; a sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:23, a sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:30, or a sequence which is identical to either SEQ ID NO:26 or SEQ ID NO:27. In the embodiments where the integrated first nucleic acid has a sequence identical to either SEQ ID NO:26 or SEQ ID NO:27, the isolated chromosome also has a third nucleic acid integrated therein, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:28, operably linked to a transcription control region, wherein the nucleic acid encodes a polypeptide at least 90%, 91%, 92%, 93%,

94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:29. In these embodiments, the isolated *Corynebacterium* species chromosome also has a second nucleic acid integrated therein, wherein said second nucleic acid encodes polypeptide which functions as a component of an amino acid biosynthesis pathway, preferably, a lysine biosynthesis pathway. Additionally in these
5 embodiments, the first nucleic acid regulates transcription of the second nucleic acid.

In another aspect of the above embodiments, a *Corynebacterium* species host cell comprising a *Corynebacterium* species chromosome as described above
10 is provided, preferably a host cell derived from *Corynebacterium glutamicum*, *Brevibacterium flavum*, and *Brevibacterium lactofermentum*. In a related aspect, a method of producing such a *Corynebacterium* species host cell is provided, which comprises transforming *Corynebacterium* species cells with a vector comprising a first nucleic acid as described above, wherein said vector facilitates
15 integration of the first nucleic acid into the chromosome of said *Corynebacterium* species cells, and selecting the host cell in certain preferred embodiments, the vector also comprises the second nucleic acid described above, and/or the third nucleic acid described above, with the first nucleic acid being physically situated to regulate transcription of the second nucleic acid. These vectors are also
20 provided.

Examples of amino acid biosynthesis pathways which may be included in the above embodiments of the invention are the enzymatic pathways that result in the synthesis of L-glycine, L-alanine, L-leucine, L-methionine, L-phenylalanine, L-tryptophan, L-lysine, L-glutamine, L-glutamic acid, L-serine, L-proline, L-valine, L-isoleucine, L-cysteine, L-tyrosine, L-histidine, L-arginine, L-asparagine, L-aspartic acid, and L-threonine.
25

In those embodiments where the second nucleic acid encodes a component of the L-lysine biosynthesis pathway, the second nucleic acid may encode one or more of the following enzymes: aspartokinase, aspartate beta-semialdehyde dehydrogenase, diaminopimelate dehydrogenase, diaminopimelate
30

decarboxylase, dihydrodipicolinate synthetase, dihydrodipicolinate reductase, or pyruvate carboxylase.

In another embodiment of the present invention, a method is provided for producing a vector: More specifically, the method comprises inserting into a vector any of the isolated polynucleotides described herein.

In another embodiment of the present invention, a vector is provided comprising any of the isolated polynucleotides described herein.

In one aspect of this embodiment, the vector further comprises a multiple cloning region into which a heterologous second nucleic acid can be inserted, thereby allowing for the regulated and/or elevated expression of a variety of second nucleic acids therein.

In another aspect of this embodiment, said vector is a shuttle vector. As used herein, the term "shuttle vector" refers to a vector that can replicate and be maintained in more than one host cell species. In a preferred aspect of this embodiment, said shuttle vector can replicate and be maintained in a *Corynebacterium* species host cell, preferably a host cell derived from *Corynebacterium glutamicum*, *Brevibacterium flavum*, and *Brevibacterium lactofermentum*, and in an *E. coli* host cell.

In another embodiment of the present invention, a method of producing a transformed *Corynebacterium* species host cell is provided. The method of this embodiment comprises introducing into *Corynebacterium* species cells, preferably *Corynebacterium glutamicum*, *Brevibacterium flavum*, or *Brevibacterium lactofermentum* cells, a vector comprising any of the isolated polynucleotides described herein.

In one aspect of this embodiment, the polynucleotide, after being introduced into the host cell, is integrated into the chromosome of the host cell. In another aspect of this embodiment, the polynucleotide, after being introduced into the host cell, is maintained as an extrachromosomal element; *i.e.*, it does not integrate into the chromosome of the host cell.

In another embodiment of the invention, a method is provided for the production of a biosynthetic product such as an amino acid, a purine nucleotide, or a heterologous polypeptide. According to this embodiment, a transformed *Corynebacterium* species host cell harboring a vector of the invention is used. More specifically, the vector comprises any isolated polynucleotide described herein.

In a preferred aspect of this embodiment, the polynucleotide comprises a first nucleic acid which regulates transcription of a second nucleic acid, where the second nucleic acid encodes a polypeptide which functions in an amino acid biosynthesis pathway, as described above. In this aspect, the first nucleic acid regulates transcription of the second nucleic acid, thereby resulting in elevated production of the amino acid by the transformed host cell.

In another aspect of this embodiment, the polynucleotide comprises a first nucleic acid which regulates transcription of a second nucleic acid, where the second nucleic acid encodes a component of a purine biosynthesis pathway. In this aspect, the first nucleic acid regulates transcription of the second nucleic acid, thereby resulting in elevated production of the purine by said transformed host cell.

In yet another aspect of this embodiment, the polynucleotide comprises a first nucleic acid which regulates transcription of a second nucleic acid, where the second nucleic acid encodes a heterologous polypeptide. In this aspect, the first nucleic acid regulates transcription of the second nucleic acid, thereby resulting in elevated production of the heterologous polypeptide by said transformed host cell.

In certain preferred aspects of this embodiment, the first nucleic acid comprises a promoter. Preferred promoters include, but are not limited to the following types of promoters: a constitutive promoter, an inducible promoter, a derepressable promoter, a heat sensitive promoter, and a cold sensitive promoter.

Where the promoter is an inducible promoter, the method for the production of a biosynthetic product as above may further comprise adding an

-12-

inducer to the culture medium. The inducer may be present throughout the growth of the host cell, or alternatively, may be added to the culture medium after the host cell has grown to an optimal density. In a preferred aspect, the host cell is self-induced, *i.e.*, the inducer is produced by said host cell at some point during its growth cycle. When the inducer is present in the culture medium, the first polynucleotide facilitates enhanced production of the biosynthetic product. Inducers which may be used in the invention include, but are not limited to acetic acid, pyruvate, ethanol, starch subunits, sugars, cellulose subunits, fatty acids, and triglycerides. Sugars that may be used as inducers in the present invention include, but are not limited to fructose, maltose, lactose, and arabinose.

Where the promoter is a derepressable promoter, the method for the production of a biosynthetic product as above may further comprise use of a culture medium which lacks a suppressor. In this aspect, the suppressor may be absent from the culture medium throughout the growth of said host cell, or the host cell may be initially cultured in a culture medium containing the suppressor, and be transferred to a culture medium lacking said suppressor when the host cell culture has reached an optimal density. In certain preferred aspects, the host cell may be self-regulating, *i.e.*, the suppressor may be removed from said culture medium through metabolic depletion by the host cell. When the suppressor is removed or depleted from the culture medium, the first polynucleotide facilitates enhanced production of the biosynthetic product. Suppressors which may be used in the invention include, but are not limited to purines, pyrimidines, amino acids such as histidine, and oxygen.

Where the promoter is a heat sensitive promoter or a cold sensitive promoter, the method for the production of a biosynthetic product as above may comprise adjustment of the temperature of the growth medium to a temperature which is greater than or less than the optimal growth temperature for an untransformed *Corynebacterium* species cell. For example, the yield of the biosynthetic product may be increased when said host cell is grown at a temperature adjusted to be greater than the optimal growth temperature for an

untransformed *Corynebacterium* species cell, or the yield of the biosynthetic product may be increased when the host cell is grown at a temperature adjusted to be less than the optimal growth temperature for an untransformed *Corynebacterium* species cell.

5 ***Detailed Description of the Preferred Embodiments***

10 The present invention relates to isolated polynucleotides that function as transcriptional regulators, in particular, promoters, in *Corynebacterium* species host cells, preferably *Corynebacterium glutamicum*, *Brevibacterium flavum*, or *Brevibacterium lactofermentum* host cells, even more preferably *C. glutamicum*.
15 host cells. These promoters are useful for regulating and enhancing the production of a variety of products in such host cells. Sources of promoters include nucleotide sequences from the 5' end of native chromosomal genes from *Corynebacterium* species, from sequences on plasmids that replicate in *Corynebacterium* species, from sequences in the genome of phage that infect
20 *Corynebacterium* species, from sequences derived from other microorganisms, e.g., *Escherichia coli*, or from sequences assembled by humans (*tac*, *trc*) which are not found in nature. Genes of ribosomal proteins, ribosomal RNAs and elongation factors show high levels of expression. The promoters of these genes are candidates for increasing expression of amino acid biosynthetic pathway genes.

25 Another reason for changing promoters of genes in biosynthetic pathways is to make the pathway independent of factors that control the pathway in the wild type organism. For example the native promoter of the operon that contains diaminopimelate decarboxylase of the lysine biosynthetic pathway of *C. glutamicum* can respond to arginine or lysine in the growth medium. Arginine increased transcription three-fold and lysine decreased transcription by one third (Oguiza, *et al.*, *J Bact.* 175:7356-7362 (1993)). Diaminopimelate decarboxylase activity decreased 60% in cells grown in minimal medium supplemented with

10mmM lysine (Cremer, *et al.*, *J Gen Microbiol.* 134:3221-3229 (1988)). Replacing the promoter of *lysA* which encodes the diaminopimelate decarboxylase is one way to make lysine biosynthesis independent of arginine and lysine levels in media.

5 Gene expression from many of the transcriptional regulators of the present invention can be easily and inexpensively regulated by manipulating the composition and/or temperature of the medium in which cells containing these promoters are cultured. Regulated expression from a promoter is especially useful for controlling the production of a gene product that is toxic to cells at high
10 levels. To allow for maximum production of a gene product, a culture of cells possessing the gene of interest under the control of an inducible promoter can be first grown to a sufficient cell density under conditions in which expression from the promoter is suppressed. Then, by manipulating the culture medium, expression from the promoter is induced. This strategy assures not only a high
15 level of gene expression in the individual cells, but also a sufficiently high number of cells in the culture, so that maximum production of the gene product by the entire cell population is achieved.

A. Definitions

20 In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided. It is also to be noted that the term "a" or "an" entity, refers to one or more of that entity; for example, "a polynucleotide," is understood to represent one or more polynucleotides. As such, the terms "a" (or "an"), "one or more," and "at least one" can be used interchangeably herein.

25 **Auxotroph.** As used herein, the term refers to a strain of microorganism requiring for growth an external source of a specific metabolite that cannot be synthesized because of an acquired genetic defect.

Amino Acid Supplement. As used herein, the term refers to an amino acid required for growth and added to minimal media to support auxotroph growth.

Chromosomal Integration. As used herein, the term refers to the insertion of an exogenous DNA fragment into the chromosome of a host organism; more particularly, the term is used to refer to homologous recombination between an exogenous DNA fragment and the appropriate region of the host cell chromosome.

Corynebacterium species. As will be understood by those skilled in the art, the terms "*Corynebacterium*" or "*Corynebacterium* species" includes those organisms previously identified in the literature as "*Brevibacterium* species," for example *Brevibacterium flavum* and *Brevibacterium lactofermentum* which have now been reclassified into the genus *Corynebacterium*. (*Int. J. Syst. Bacteriol.* 41: 255 (1981)). Accordingly, the term "*Corynebacterium* species" is used herein interchangeably with "*Corynebacterium* species and *Brevibacterium* species."

Enhancers. As used herein, the term refers to a DNA sequence which can stimulate promoter activity and may be an endogenous element or a heterologous element inserted to enhance the level, i.e., strength of a promoter.

Inducer. As used herein, the term "inducer" refers to molecule which acts to stimulate transcription from an inducible promoter. The inducer may be produced by a host cell, or added to a culture medium in which the host cell is being grown.

Isolated Polynucleotide. As used herein, the term is intended to mean a polynucleotide, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. However, a nucleic acid molecule contained in a clone that is a member of a mixed clone library (e.g., a genomic or cDNA library) and

that has not been isolated from other clones of the library (e.g., in the form of a homogeneous solution containing the clone without other members of the library) or a chromosome isolated or removed from a cell or a cell lysate, is not "isolated" for the purposes of this invention. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Lysine Biosynthetic Pathway Protein. As used herein, the term "lysine biosynthetic pathway protein" is meant to include those peptides, polypeptides or proteins, and enzymes, which are directly involved in the synthesis of lysine from aspartate. These proteins may be identical to those which naturally occur within a host cell and are involved in the synthesis of lysine within that host cell. Alternatively, there may be modifications or mutations of such proteins, for example, the proteins may contain modifications or mutations which do not significantly affect the biological activity of the protein. For example, the natural protein may be modified by mutagenesis or by introducing or substituting one or more amino acids, preferably by conservative amino acid substitution, or by removing nonessential regions of the protein. Such modifications are readily performed by standard techniques. Alternatively, lysine biosynthetic proteins may be heterologous to the particular host cell. Such proteins may be from any organism having genes encoding proteins having the same, or similar, biosynthetic roles.

Mutagenesis. As used herein, the term refers to a process whereby a mutation is generated in DNA. With "random" mutagenesis, the exact site of mutation is not predictable, occurring anywhere in the genome of the microorganism, and the mutation is brought about as a result of physical damage caused by agents such as radiation or chemical treatment. rDNA mutagenesis is directed to a cloned DNA of interest, and it may be random or site-directed.

Mutation. As used herein, the term refers to a one or more base pair changes, insertion or deletion in the nucleotide sequence of interest.

Operably Associated. As used herein, the term "operably associated" refers to a association of nucleic acid elements in a functional relationship. A nucleic acid is "operably associated" when it is placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably associated with a polypeptide coding region if it affects the transcription of the polypeptide coding region. Operably associated means that the nucleic acids being associated are typically close together or contiguous and, where necessary, join two polypeptide coding regions, contiguous and in reading frame. However, since enhancers generally function when separated from the promoter by several kilobases, some polynucleotide elements may be operably associated but not close together or contiguous.

Operon. As used herein, the term refers to a contiguous portion of a transcriptional complex in which two or more open reading frames encoding polypeptides are transcribed as a multi-cistronic messenger RNA, controlled by a cis-acting promoter and other cis-acting sequences necessary for efficient transcription, as well as additional cis acting sequences important for efficient transcription and translation (*e.g.*, mRNA stability controlling regions and transcription termination regions).

Parental Strain. As used herein, the term refers to a strain of host cell subjected to some form of treatment to yield the host cell of the invention.

Phenotype. As used herein, the term refers to observable physical characteristics dependent upon the genetic constitution of a host cell.

Promoter. As used herein, the term "promoter" has its art-recognized meaning, denoting a portion of a gene containing DNA sequences that provide for the binding of RNA polymerase and initiation of transcription and thus refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes, upstream of one or more open reading frames encoding polypeptides. Sequence elements within promoters that function in the initiation of transcription are often characterized by consensus nucleotide

sequences. The promoter sequence includes proximal and more distal upstream elements. Examples of proximal elements in bacterial promoters include a -10 region and a -35 region, which are discussed in more detail, *infra*. Examples of more distal elements include operator regions and enhancer regions. As used
5 herein, the term "endogenous promoter" refers to a promoter sequence which is a naturally occurring promoter sequence in that host microorganism. The term "heterologous promoter" refers to a promoter sequence which is a non-naturally occurring promoter sequence in that host microorganism. The non-naturally occurring promoter sequence may be from any prokaryotic or eukaryotic
10 organism. A synthetic promoter is a nucleotide sequence, having promoter activity, and not found naturally occurring in nature.

Promoters may be derived in their entirety from a native gene, or be hybrid promoters. Hybrid promoters are composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA
15 segments. Hybrid promoters may be constitutive, inducible or environmentally responsive.

Useful promoters include constitutive and inducible promoters. Many such promoter sequences are known in the art. *See*, for example, U.S. Pat. Nos. 4,980,285; 5,631,150; 5,707,828; 5,759,828; 5,888,783; 5,919, 670, and,
20 Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Press (1989). Other useful promoters include promoters which are neither constitutive nor responsive to a specific (or known) inducer molecule. Such promoters may include those that respond to developmental cues (such as growth phase of the culture), or environmental cues (such as pH, osmoticum, heat,
25 dissolved gases, or cell density).

Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, reduced temperature, or the presence of light. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different cell types,
30 or in response to different environmental conditions. Promoters which cause a

gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical or similar promoter activity.

5 **Relative Growth.** As used herein, the term refers to a measurement providing an assessment of growth by directly comparing growth of a parental strain with that of a progeny strain over a defined time period and with a defined medium.

10 **Stringent Hybridization Conditions.** As used herein, the term "stringent hybridization conditions" is intended to mean overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (750 mM NaCl, 75mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

15 **Suppressor or Repressor.** As used herein, the terms "suppressor" or "repressor" refer to molecules which act to block or reduce transcription from an derepressable promoter. The suppressor or repressor may be produced by a host cell, or added to a culture medium in which the host cell is being grown. Furthermore, during the growth of a host cell, the suppressor substance can be
20 metabolized by the host cell, thereby removing it from the culture medium, thereby increasing transcription from the derepressable promoter.

Transcription factor. As used herein, the term "transcription factor" refers to RNA polymerases, and other proteins that interact with DNA in a sequence-specific manner and exert transcriptional regulatory effects.
25 Transcriptional factors may be transcription inhibitory proteins or transcription activator proteins. In the context of the present invention, binding sites for transcription factors (or transcription complexes) are often included in the transcriptional regulatory element(s).

Transcription factor recognition site. As used herein, a "transcription
30 factor recognition site" and a "transcription factor binding site" refer to a

polynucleotide sequence(s) or sequence motif(s) which are identified as being sites for the sequence-specific interaction of one or more transcription factors, frequently taking the form of direct protein-DNA binding. Transcription factor recognitions sites which bind transcription inhibitory proteins or transcription activator proteins are also referred to as "operator regions." Typically, transcription factor binding sites can be identified by DNA footprinting, gel mobility shift assays, and the like, and/or can be predicted on the basis of known consensus sequence motifs, or by other methods known to those of skill in the art.

Transcriptional Complex. As used herein, the term "transcriptional unit" or "transcriptional complex" refers to a polynucleotide that comprises one or more coding regions for polypeptides, a cis-acting linked promoter and other cis-acting sequences necessary for efficient transcription of the structural sequences, distal regulatory elements necessary for appropriate transcription of the structural sequences, and additional cis sequences important for efficient transcription and translation (e.g., mRNA stability controlling regions and transcription termination regions). In bacteria, a transcriptional complex may comprise a single coding region, or one or more coding regions, e.g., as part of an operon.

Transcriptional Regulatory Element. As used herein, the term "transcriptional regulatory element" refers to a DNA region which activates transcription alone or in combination with one or more other DNA regions. A transcriptional regulatory element can, for example, comprise a promoter, response element, negative regulatory element, silencer element, gene suppressor, transcription terminator, and/or enhancer.

B. Microbiological and Recombinant DNA Methodologies

The invention as provided herein utilizes some methods and techniques that are known to those skilled in the arts of microbiology and recombinant DNA technologies. Methods and techniques for the growth of bacterial cells, the

introduction of isolated DNA molecules into host cells, and the isolation, cloning and sequencing of isolated nucleic acid molecules, etc., are a few examples of such methods and techniques. These methods and techniques are described in many standard laboratory manuals, such as Davis *et al.*, *Basic Methods In Molecular Biology* (1986), J.H. Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1972); J.H. Miller, *A Short Course in Bacterial Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1992); M. Singer and P. Berg, *Genes & Genomes*, University Science Books, Mill Valley, California (1991); J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989); P.B. Kaufman *et al.*, *Handbook of Molecular and Cellular Methods in Biology and Medicine*, CRC Press, Boca Raton, Florida (1995); *Methods in Plant Molecular Biology and Biotechnology*, B.R. Glick and J.E. Thompson, eds., CRC Press, Boca Raton, Florida (1993); and P.F. Smith-Keary, *Molecular Genetics of Escherichia coli*, The Guilford Press, New York, NY (1989), all of which are incorporated herein by reference in their entirety.

Unless otherwise indicated, all nucleotide sequences newly described herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc.). Therefore, as is known in the art, for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art.

C. Polynucleotides

Certain embodiments of the present invention are directed to an isolated polynucleotide comprising a first nucleic acid, the sequence of which is related to, or identical to, a nucleotide sequence, or fragment thereof, selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, and SEQ ID NO:30. SEQ ID NOs 1-3 have been previously described. *See, e.g.,* Reinscheid *et al., Microbiology* 145:503 (1999); U.S. Patent No. 5,700,661; and U.S. Patent No. 5,965,391. SEQ ID NOs 23, 24, 26, 27, 28, 30 and 33-35 have been previously described. *See, e.g.,* Ben-Samoun *et al., FEMS Microbiology Letters* 174:125-130 (1999) Brosius, J. *et al., J. Biol. Chem.* 260, 3539-3541 (1985), and Amann, E., *et al., Gene* 69: 301-315 (1988), Moeckel, *et al., J.Bacteriol.* 174:8065-8072 (1992), Keilhauer, *et al. J.Bacteriol* 175:5595-5603 (1993), Patek *et al., Appl. Env. Microbiol.*, 60:133-140 (1994). Isolation of polynucleotides comprising nucleic acids the sequence of which are identical to SEQ ID NOs 4-22 are described herein. While not being bound by theory, it is believed that polynucleotides having these sequences can be used to regulate gene expression in *Corynebacterium* species. The genes regulated by SEQ ID NOs 4-22, determined by similarity of the genes to genes identified in other organisms, are listed in Table 1A, along with putative exogenous regulatory molecules. SEQ ID NOs 1-22 are listed in Table 1B.

TABLE 1A
Nucleotide sequences that can be used to regulate gene expression

Seq. I.D. NO:	Gene*	Regulatory Molecule†
1	pta	acetate
2	aceA	acetate
3	aceB	acetate
4	adh	ethanol
5	aldB	ethanol
6	poxB	pyruvate
7	ldh	pyruvate
8	amyE	carbon
9	malZ	carbon
10	bglX	carbon
11	gam	carbon
12	glgX	carbon
13	hisD	histidine
14	pyrR	pyrimidine
15	purD	purine
16	hrcA	temperature
17	htpX	temperature
18	dnaK	temperature
19	ctc	temperature
20	grpE	temperature
21	clpB	temperature
22	narA	oxygen

Sequence I.D. NOs 1, 2, and 3 have been previously described. The remaining sequences were discovered in ADM's *Corynebacterium glutamicum* genome sequencing project.

* Putative genes regulated by sequence I.D. NOs 4-22 were determined by homology to genes identified in other organisms, e.g., *Escherichia coli* or *Bacillus subtilis*.

† Putative regulatory molecules associated with the regulatory regions of SEQ I.D. NOs 4-22 were determined by analogy to regulatory regions identified in other organisms.

Table 1B**Nucleotide sequences that can be used to regulate gene expression****SEQ ID NO: 1**

CACCGACAACGGCAACACGCAAAGGGCGAGACATATAAAGTTCGATTCCCTAAAGGGGTTCTAAAAAATG
 TGGAGTATGTGAGCGGGGTTCCTAGTAGATTGCGACTCCTATCGGGGTGCGACTGCTAATGGTGCCCTG
 CTATCAACCCTCCATGATACGTGGTAAGTGCAGACTAATAAAGGCCAGTCGGGGAGGATTGGGGGCTTTG
 CTGGGGGCGAGTTTGTACGCTGCGCGCTTTTCATAGACCCCATTAATGTGGGGTGAAGAGCTGTAAAGTA
 CCGCTAAAACTTTGCAAAGGGTGCTTCGCAACTTGTAAACCGCTCCGTATTGTTTTCTACGGCAATAAGC
 ATTTGTGCTGCTCAAAGCGTGGAATTGAGATCGGTTTGAAAATTACAAAATAAACTTTGCAAACCGGGC
 TGTACGCAAGGCGGACGAACGCTAAACTATGTAAGAAATCACAACTTCCCTCAGTAGTGCCAGGAGGCA
 CAAGCCTGAA

SEQ ID NO:2

ACTCTTTTAAGAAAAGCACTCTGACTACCTCTGGAATCTAGGTGCCACTCTTCTTTTCGATTTCAACCCTT
 ATCGTGTGGCGATGTGATCAGACTAAGTATCACCGTCACCAGCAAAGGGGTTGCGAACTTTACTA
 AGTCATTACCCCGCCTAACCCGACTTTTATCTAGGTACACCTTCAAACCTACGGAACGTTGCGGTG
 CCTGCATTTTCCATTTTCAAGCATTTGCCAGTACATCTGTACTAGCAACTCCCCGCCCCTTTTCT
 GCGAAGCCAGAACTTTGCAAACTTCAACAACAGGGGTGACCACCCCGCACAAACTTAAAAACCCAAAC
 GATTGACGCACCAATGCCCGATGGAGCAATGTGTGAACACGCCACCAGCAAACCGATGCACATCACGT
 CGAAACAGTGACAGTGCACTAGCTCATCTTTGTGGTTGGCACCGCCCATTCGCAATCAGCACTTAAGGA
 AGTGACTTTG

SEQ ID NO:3

TTGCGTGGTGGCGTGGTTCACACATTGCTCCATCGGGCATTGGTGCGTCAATCGGTTTGGGTTTAAAGT
 TTTGTGCGGGGGTGGTCACCCCTGTTGTGAAGTTTGCAAAGTTCTGGCTTCGAGAAAAAGTGGGCGGGG
 GAGTTGCTAGTACAGATGTACTGGGCAAATGCTCTGAAATGGGAAATGCAGGCACCGCAACGTTCCGTA
 GGTTTTGAAGGTGTGACCTAGATAAAAGTCGGGGTTAGGCGGGGGTAATGACTTAGTAAAGTTCGCAAC
 CCCTTTTGTGCTGGTGACGGTGATCACCTAGTCTGATCACATCGCCAAACACGATAAGGGTTGAAATCGAAA
 GAAGAGTGGCACCTAGATTCCAGAGGTAGTCAGAGTGCTTTTCTTAAAGAGTTTTCACAACCGTTAACG
 GCGTAGCCAAACAAGAAGGATTGCGATTCTTCTGGTTTAGGCACAGGTCATCTAAACCCATGCTTTAAAA
 GGAGCCTTCA

SEQ ID NO:4

ATTGGAATGGAGATTAGAAGAATCCTGGGAATGTTGGTGTGTGTTGCATGTCTGTTGAGACTATCTAG
 TAGATGCGGCTGCGCTCCTTAATTGCATGCTGGGGTGGTGGGGAATGGGTGGTTGGGGGCGTCGAAAAGC
 ATTTTGGTGCTTCTAAGGGAATTGTGTGAATCTTGGAAAGCTAATTGAAAAACATTCCCATTAGTGGGT
 GATTTGCTGGAGTTTGTGAATCTATTTTTCGAAATTTTTCAGCGTGCGGGGTGGTTGTTTTTTTACAAT
 TGCCAGTTCATTCACGGTTGTTGAAATGTTTCGGGGGTAAATAACTCACTTTCTATTTTACCTTGTGGG
 ATTTTCGTAGGGTGGACGATGGCAGCAATTGAATGTTGTAAATCACACAATTGCAAGGATTGTAATTTAA
 GGCACATCTATGTCGGTGTGAAATTACATGTGCCAGAAGAGCAATTGCCAAGTAATCCAAGCGAGAAGG
 AGTGAGTTT

SEQ ID NO:5

-25-

CGATCATCGAACTCGGCGAAGAAAACCGAAACCTCAAAGAATCCCTGCGTAAGGTCACAGCTGAGAATGA
GCAGCTCAAAGATCAATTACGCAGCGGGCGTCCGCGTGCGGAGCTGGTGCACGTGCCCCGCTCCACCGCG
GTGGTCAATGTGGGAACGCCGCAAGGGGCGTTCCAAGTAAAAACATGCTTGTGACGCCGCTTTCTAGCAA
ATTAAGCGGGCACCTCCATTTATCTTTTGGAGGTGCCCGTTTCGTGCTTTCGCCAATTAGATACATGCAT
AACCACCCGAACAGGGGTAAATAACTTTTGAAAGGCTTTCGCGCTTGAGCTGCGAGAATTTTGAGAAAAGG
GGGTGAATTTAACAGGGGTCTAGCGCGGATTGATTTTCGTGAATATGGTGGCTGCTAAGCGTGCGAATG
TGC GCGTTATCACAATCGTTGACCAAGTGTCACCTGACGCACAGGTAGTGCTCAGGTGGAGGTGGCCCCAA
AGGAGACCCA

SEQ ID NO:6

TTTGTAGACCACGGCGCTGTGTGGGGATTTAAGACGTGCGAAATTGTAGGGGACTGTCAGTGTGAGTCCGG
TTCTTTGAGGGCGCTTAGAGGCGATTCTGTGAGGTCACTTTTGTGGGGTCGGGGTCTAAATTTGGCCAGT
TTTCGAGGCGACCCAGACAGGCGTGACAAGATTGACTAAAAAACCGAAGTTTGGCACGTGTGTTTGGTTT
CTCGGGGTCTAAACCGGACAGGCGTGACAAGATCTGGCGAAATCGCAGGTTTGTGTACGCTGTCTGGT
TTTACCTTTTGGGGGCCGAACTGCGCTGAACTACTCGGATCGACCAAGCAGTTTGGCTCTCCAGCGCTCT
GATCAAGCACCAACCGCTCTAAATCACCAAGGCACTCGTAAACCCGTGGCAGATAGAGAAAGTGT
GGCAGCAACTCGAATTGAAGAGCACAATTGAAGTCGCACCAAGTTAGGCAACACAATAGCCATAACGTTG
AGGAGTTCAG

SEQ ID NO:7

AAAAACAGCCAGGTTAGCGGCTGTAACCCACCACGTTTCGGCAACAATGACGGCGAGAGAGCCACCACA
TTGCGATTTCCGCTCCGATAAAGCCAGCGCCCATATTTGCAGGGAGGATTCGCCTGCGGTTTGGCGACAT
TCGGATCCCGGAACCAGCTCTGCAATGACCTGCGCGCCGAGGGAAGCGAGGTGGGTGGCAGGTTTGTAGT
GCGGGTTTAAAGCGTTGCCAGGCGAGTGGTGAGCAAGACGCTAGTCTGGGGAGCGAAACCATATTGAGTC
ATCTTGGCAGAGCATGCACAATTCTGCAGGGCATAGATTGGTTTGTCTGATTTTACAATGTGATTTTTTC
AACAAAAATAACACTTGGTCTGACCACATTTTCGGACATAATCGGGCATAATTAAAGGTGTAACAAAGGA
ATCCGGGCACAAGCTCTTGTGATTTTCTGAGCTGCTTTGTGGGTTGTCCGTTAGGGAAATCAGGAAGT
GGGATCGAAA

SEQ ID NO:8

CCAGATCTCGGGATGCTTCGTAGATCAGCATGGCGCCGAGGTGCTTGGCTGCTGCACGGAAAGCTGCGTT
GTCGCTGCGCTCGTCGCGCAACAGGGGTAGGCGGCTAGCAACGAGTGGGTGGTTGACGATGGTGATGTCC
ATAGCTTACATGTTAAATCATTGCCGCCCAGAGAAGACCGCGCGGGCGAATTTGGGCTTGGAGGGAAC
CAAACGGCCACTTTTCCAGTCCAACAAAGTATGAGGATTAATTTGCCCACTCCAAGAGCTCGCCACG
AGCTGTGTTTGTGCCACCCCTGCTGTGCCCGCGCTTCCCACTGATTCTGGCGCGCAGTTTGTATCCA
CCAGGCACTGTCCGCTCTCTTGCACCTATGCCGCAACCTCACCTTGTGTCCACACCGCCGAGAAT
TTAGGAAACCGCGCTCTGACCGGCTCGTGAAATCGAAGACACCGACCAACTCGCACACGCATTGG
AGCGCCTGAC

SEQ ID NO:9

-26-

GGGCATCATGGGCATTGTTGGTGGATTACCATGCCTGTCGCAATCGCTAGGACCAGGGATAAGAACCTC
GTGTGGTTCCCGTGGTCTTTGGTGCATCGATGTTCTCGGTTATGTGGGAACGTGGCTGTGGCCGTCCC
AAGGCTGGTACCTGTGGTCATTCTTCTTGGTTTAGGTGGACTCTGCTTCCCGATGGCTATCGCCCTGAT
TCCAGCGCGTACGAAAGATCCGAGAATTACCGCAAGCTTGTCTGGATTGTGTCAGCCGGTGGGTACATT
CTTGCAGCCCCTGGGCCATTGGCAGTGGGAGCGATCTACCAGGCGATTGGCTCCTGGTCAGAGATCCTCG
TTGGTTTGGCCTTGGGCACAATAGTGTGTGATTTGTGGGATTGAGAGCAGCACGCAATGTGACGGTTGA
TGATGAATTGAGGAGATCAAAGTAGCCTCAACTAAGCGTCGCGATAAGAACGAGGGGCAAGGCTGATGTA
CTCTGTCAACCATGGATAAACCGGTCGTGAGGGATGCAGCTCTGCTGATTTTTTCGCGCTGTGCTCGGAGT
GATCTTTGTGGCACACGGGTGGGAAAAGCTGTTTCTCTCCGGAGTTACCAAGACAACAGGACAATTTTCA
GCCTGGGGAGTGCCTCAACCAAGCTCTCGGTGTGGATCACATCGATCTCTGAGCTGCTCGGTGGTGCCT
TCCTAGTGGTTGGTTTGCTCACCACCTTTGTTGCTGGTGCATTAGCGCTGTTGATCGCCGCTGCTATTTA
CTTTGTGCACTTGAGTTCCGGCTTTTTTACAGTTGATAACGGCATCGAATTCCCCTTGCTCATCATTTGTT
TCTTTGCTCGTGATCGTTGTGTTTGGTTCTGGTAGAGCCAGCGTTGATGGGGTGCTCACGCGTGGTTGAC
TGTAAGTGCATTAAGCCGCGCTGTCCGCCAAATTAGATGGTGAGCCGACAGGCTTGGATGATGCAGTAA
TTGAAGCGCACCTCGCTAATTGTGAAGAGTGCAGAAATTACTACAACCGTGTGCTGAGTTGAATCGGAT
GCTCAATTTTTGCGCCGCGGAACTCGCACCCGTGACCCCGCTGATCTATCAGAGATCATCTGGCAGAG
GTGGAACCCAGAATGGCGCAGGCATGCCAACGCCAAGGTTGTGGGATCCCTGCTATCGCGAGTGTGTTGG
TGATCCTGGGTGTGGTTTACCTCGCCTGGGGTATCACAAATGTGGGGGATTTCGGCGTCGATAAGCGTCCA
AGAAGACCCGCTCACCTCGCGCTGTGCGGAGGCCGTTGCCTACCGCATTGCTCTTTCTGTGGGGCTG
TTATTTGCGGCGTGGAAGCCGCGGATTATCGCGGGCATGCTCCCGATTTTGGAAACGTTGTGGACATTTA
TGCTGTTTTGCTGCGCGCGATCTCGTGTGCGGTGCGCGATTACAGACGGGACTGTCCATTGGTCT
GCTGTTGATTTTACGATTGTGCTGTGCTTTGCGTTGGTGAATAGTTCTGGACCGGGTATTTTTCGCGCG
ACATGGAACCTCATGAACGCCGCGCCCGCTAAGTGGGAGGC

SEQ ID NO:10

TAATACGGCATCCGGCCTGAAAGGACAGTGTGGAACCCAAAAATCATGCCCTCGCAGAATCGTGTTTAAG
GGGTTAAAAACGCTCGACCCACACTCTGACCCATCCGTGAACTAGACCTCTTAAACGACTTCTGGTTG
GTGAGGTGTGAAACCGCACTGCGGGGCCCAATTACAAAAATCCATACGAGTGTGCACACCCGATTTTC
ACATCGCTTCGAGACCTCCCTTTTTGACACCTTAATTGTCTAACCCCGTATAGGTGAGAAATGTTGGACA
AGTGTCTGTTTTGTGGGGGAATCTGACTACGATGGTAAGAAACAGGGAAAGGGTTACCATTTATGTCT
CAAGAGTAGCCTCAAATCGGCTCCCGCTCTCTCGTGTCTTGAACAAGACGGCTACAATTCGCGGATC
TCGACGGCGACGGGTACTTGCACCTTATGAAGATTGGCGTCTAACCCAGCAGAGCGTGCCGTGACCT
GGTGAAACGA

SEQ ID NO:11

CACCGCGTCGCGGATCTCTCCCTTCCAGACTCAGAAGTTGCTTCGGCAATTGCATTGAGTGGAAACCAG
GCGTCGGATTCACTACTTGAGCTGCTGATTTGTAGGTTTTAAGACCTTGAAGATATAGTTAATTCTCGTT
GCAAGGAACCAGATTCCAAGCAATGCGTATTCTCCATAGCTCAGTTGGCAGAGCATTCGACTGTTAATC
GAAGGGTCACTGGTTCGAGCCAGTTGGAGGAGCAAATTGAAACCCACTGTTTTTAAACAGTGGGTTTTT
TTGCATGTTTATACAGTTAACGAACCCCAATTTGTAACTTCCACTCTCCTAGCTATGATGAATACTCG
TTGCAAGGAAGTAATTCTTTCCAATACTTATTCTCCATAGCTCAGTTGGCAGAGCATTCGACTGTTAA
TCGAAGGGTCACTGGTTCGAGCCAGTTGGAGGAGCAATACACACAGCCCGCGTTTTTCTTAAACCGG
GGTTTTTGT

SEQ ID NO:12

AGTGACAGCTCGCGCCGATCGTTGATGGCAAACATGTCCAAATTAAGGCGCATGCGACCAAGGATGGTG
AGCATTTTGCCCTGATTCTGGCTGAAGTACATCGAAATGCGGTGATCTTGCCACGGCAGCATGATGCGCAT
CTTCTGAGCTCAAATAGGTGGTAGCCCAAGGAATCAACAGCCTCCGTCACTCGATTCAAGTCAACTGGGAA
AGGAATGGAGGTGTTGGGTGACGGGACGTTTCGATCATTCACAGAGGTGAATCCATTCTCTCCAGTATC
TCAAAGGTGAAAGCGGGTTTAATTCAAGGTAATCTGGGGTGGTCATTTTAAGTTTAAAGTCTAATTCAA
TGAATCTGTGATGACCCAAATCAGAAACTTGTACGTGGGGAATACAATAGGTAAATATCGGGGCTTAAG
AACTTGTGTGAGGCCGCTTGGATTGGGCACCGAGCTCGAAGAATTCGATTCAACCTTTTAAGGGAGA
ACTTTTCGCC

CGTGGCGTGCCACCCATTTAAGTCCCgcggggagactGAAGATGGTGGTGCCGTGCGGcGAGGCGTTTCTGC
CTGCGTAATGGGGTTAAGTGGGATGAACGGGGGgAGTCAGACGTGCGACAGCGCCCTTcCGGGTATGCCAA
TCCCAGACCAATTTCTCGGGGAAAAGGAATAAAATGGCTTGTGGTCAGACTCAACGGGGCTTCTCCAAGTC
AGTGGATTTATGAGGTCAGGTCCGATGGGTGGGTACACACCGGGTGTCTCTACAACGATCAATTGTACAGATTGCAT
GGCATGCTGTACCATCTGCTTTAAGCATTTTGGTGTTCACtGTTGTTAACAGTGTTCACCGTGGAGCA
CTACCTTAGATCATAGTCAGCATCTTGGGGTGAATGTGACACCGTACGCTATAGTGTcAGACAACAACCA
GGAACCTGGTCGTTGCAGAGCTTTTTGCAAAATTGGACATCCTTTAAcGGACCGCACAGAGAGCGGGGAA
GGAGGTCAcG

GCCATGGCGTTGCGGAAATCGTAATCGGCCATTTTGTGCGTGGGGAAGAGGTGGATGTGAGTGTGGGGAA
 CATCGAATCCTGCGATGATGTAAACCATCGAGGGGCGTCGAATGCTGTGCGGATTCATTTCCGATGAG
 CTGGGAGGCCCTCGTTTACTTCGCTCCAGATGTTCTGAGGAAGTTCGGTCACGCGTCAACTCTTCGCAACG
 GGTACGACTAGGTTGTGCCCTAGGTGAGGGGTTTCAGTGGATAGAAAAGCCACGACATTTCTCGGAACGAT
 ACACAAATCGGCCGGGGAGCTCGCCATTAAATAATTTTCGTGAATACAGAAGCCATATGCACAGACTACTA
 CTTGGCGTGCAACCAAAATTGAGGTTGCATAAAATTAATGCAGGTGAGCCGGTCATTTTAAAGCGCTTTTCG
 ACGCCACTTTCAACCATTTCCGAACCGCCAAGAATACTGGAAATAGCTTGGATCAAGTTTTGCAGGATAAA
 CTGTGCAACC

-28-

ACGCTGATCGTTTTGAGCTTATCGACGCTCGCCTGCGCTCAGCTGGTTTTGATTTGGTACGAGGTATCCAA
CTGGGCGAAACCCGGCGGAGAAATGCAAGCACAAACATGGGCTATTGGGTCGACGGCGACTGGTGGGGTGCT
GGCCCGGGCGCGCACTCGCACATCGGCGACCGCGCTTCTACAACATCAAGCACCCAGCGCGTTACTCCG
CGCAGATTGCGGCCGGCGAGCTGCCATTAAAGGAAACAGAGCGGCTGACGGCGGAAGATCACCACACCGA
GCGCGTCATGCTTGGTTTGCGCCTGAAACAAGGCGTGCCGCTGAACCTTTTCGCACCCGACGCGCGCCG
GTCAATCGACCGTCATATCGCAGGAGGCTGCTGCACGTCAATGCGCTGGGCAACCTGGCGGTGACCGATG
CGGACGTTTTGCTTGCCGACGGCATCATCGCCGACATTTTGCTTAGTGAAGAAGACTAAATATTTAGTAG
GTTACAGAC

SEQ ID NO:17

TGGCGACTATGCCACATAGTCGACTACCTTGCATAGTTGACTATTTGATTGAGTTGAACTGTGTGAGTGT
ATGAAGAGAAAAATGAAGAGAAAAACAGCTGCACATGGTTTCAAAGATAAGGAAGTGAAGCATGAGCAT
CGAGCCAGGAATCCCCACGCTTGGACCGCTTGAGAAACAAGTCATGCACATTCTGTGGGATCAGGAAAA
TTGACAGTCCGTGAAGTCATCGAATTCCTTCCAGGTGATCCTGCGTACACAACGATCGCAACCGTCCTGC
GTCATTGGGCGAGAAAAGGCATGGTCACCATTTGTAAAAGATGGTCGGACTGCTCGACACAGCGCGTTGAT
GAACAGGGAAGAATACACCGCTGGCGTCATGGATCAGGTGCTGTGACCGAGTCGGGATCGCACCGCATCA
ATTTGTCATTTGCTGGATACGATCACGGCGACTGATCGCGAGCTGCTTCTGGAGTATCTGCAACAGCAGG
AGGGCAGGAA

SEQ ID NO:18

AGACTTGGCGTCAAAGTTGAGTGGGACCTGTTGAAAACAGCTCGTATCTTCCCTGATTGGGTGGTTGAAA
TTAGGGGTAAACCCGGATTTTTTCTCAAGTGAAGCGCTTTGACCTGTGTAAATTAAAGAAAGTTGAGTCTA
GTGGGAACAACCTTTGTGGCATTACCGTTGCCATATATGTAAGCTTGAGTCAGGCAGGCTCAATGAGGAG
TTTTTCTTACCGGCGAAAGTCGGTGAAGCAAGTCAAAGCTCAAGCCGTGGACAATACTAAAATCACCTA
AAACAGGAGGCACCAT

SEQ ID NO:19

CGGGAGCCGTTCTTCTCCACGATGGGGCGCTCGGTGAAATCACTGAAACGCGCGGTTTCCAGAAGCATCA
CCACGATGGCGTCACTGACCTCAAAGATGCAGTGTGCTCATCACTGAAGACGGTGTTTTCTTTGAAGCC
AAGACCTGCATAAAAAACGCTTAGATGCTGCAAGGTGAGTACTGGTAGGTTGATGAAAATCATGTCGTG
CTGAAGTCGTGCCATGGGAGTTCCTTCCCTCAAAGTGCCTTTTGGCTAATGTGACCCCAACCGATATGC
TGCCGGTTTTTGGATTAGGTTTGGCCATCGTGCTAGCATATTTAGGTCTCGGCGAGGGTCAAGTACTTTT
AGTGCTCAACCGTTATCGACGCGATCTAGACTTCTAAAGTGCATTTTGTGCGCTGCCTGCGAAGACTCG
ACCAAGACATTCGAGTCGTCGCGGCATTTTTTATTTTCGCGGCGAGTGTCACCTTCATCCATGAGG
AGAAATCACT

SEQ ID NO:20

TGGATTGTCTGGCGGACAAAGTTCGCACCGATAAACCCGGCACCCGGTCACAAGCAAAGAAGTCATGG
TCGTATAGTAGTTGCGGCTCCTGCAGTTGCGCGTGTATCGCTGACGGGCATGTCAATTACGCCGTGAC
CCAGCTTCGGTGCGGTGTGAAAACGTATGGGGAAGCCACGGATAGTGCATGCCCACTTCTCCACATGG
GCTGTGCTCAGCTGTGCTAGGGGCCACGGCTTGCGGACTTCTGAACACCCGGACGAAAGCCTCCATTAT
TGATGGCACCCACCCGGGGTGGTATCGGTGCTAAGTTCACAGTGGTCGGAGAGAAAGTCAAGTCAAGT
CTGATTACATGTACGAAATCAGGGCGGTCAAATGCCAGTTCCACCCGGCTTCACTAAGGCGGGACGACC
ATAGAAACGCAGCTTCTGACATTGAATGCGCCGGTTTGCTATGGACTACACCGCATATGAGGAAAGGGCT
TGAAACGCAC

SEQ ID NO:21

CTTAATGAGTACGGCGTGGCTTATATCGGTTTCGCTCCGTTTGAAATCGGCATGACCATCATCGCGCCGA
 TCGCGGTGCTCGCAGGCTTTACTATGGGGTTGGGTGGGCGTCGTTGATTGTTGCGATCGTGATTTTGG
 CCTCGCGTGGGGTCTGAAGTGGTTGCCGGAGCGCGGACATGTCGCGGCGGAGGGTAAGCCGCAATAAAGG
 TTGGAAGCGCCGGGTCTAGGTCCGGCGCTTCTTTCGTACGCTTTTCGACGCCTCCCTCCACGTAATATTA
 AAGTTACGGGTTTTCCCTGATGCTTAAGTGGTAGTCAGTGCTTAAACTTGACTGCGGTCCACTCAATTTA
 TTTTCAAATTTTGAAGTTGAGTGGAACTACTCAACTCTTTGTGCGTTATAGATATTAGAGAGTTAAA
 TAATGGCGCTTGACCTGCAGGAAATTGAGATCAACACTGATTGTGTAGGTTGGCGCCCAACAAAGAAAGG
 GCGTTGAAAG

SEQ ID NO:22

TGGCCATTCTGTGGTCCATGCTCGCCCTGTTCTTCTTCACTGGACTGGGCAACGCCGCACATTCAAAC
 AAATGCCCATGATTTTGCCCAAACGCCAAGCAGGTGGCGTGATCGGCTGGACCGGTGCCATTGGTGCCTT
 CGGCCCCCTTCATTGTGCGGTGCTTGCTCTCCTTCACTCCAACTGTCGCGTTCTTCTGGGGCTGCGTGGTG
 TTCTTCATCATCGCCACCGCTTTGACCTGGATCTACTACGCCCGCCCGAACGCTCCATTCCCGGGATAAA
 CCGAAAGGCCAATCCATGACTACAATACTTCTTCTGGGAAGTCTTCTGAACAGTCTTCTGAAAAGATCA
 ACCCCCTCTTCAAGCTCGGCAGTTTCCCTAAGAAAAGGCACCGTCGTTCTGAAGGCCAGCAGATTTTCTT
 TCAGGGCGGACGCCAAGCCGATGTGTTTTATCGCAACCCGATGGGCGTTCGATAAAAGTCGTGCGCTCC
 CACACATGGC

In certain embodiments, polynucleotides of the invention comprise a first nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or a complementary sequence thereof. Of course, this embodiment also encompasses a first nucleic acid, the sequence of which is identical to any of the recited nucleic acid sequences.

By a polynucleotide comprising a first nucleic acid, the sequence of which is at least, for example, 95% "identical" to a reference nucleotide sequence is intended that the first nucleic acid sequence is identical to the reference sequence except that the first nucleic acid sequence may include up to five mismatches per each 100 nucleotides of the reference nucleic acid sequence. In other words, to obtain a first nucleic acid, the sequence of which is at least 95% identical to a reference nucleic acid sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The reference (query) sequence may be any

one of the entire nucleotide sequences shown in SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, or SEQ ID NO:22, or any fragment of any of these sequences, as described *infra*.

As a practical matter, whether any particular nucleic acid sequence is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to, for instance, a nucleotide sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or a complementary sequence thereof, can be determined conventionally using sequence analysis computer programs such as a OMIGA® Version 2.0 for Windows, available from Oxford Molecular, Ltd. (Oxford, U.K.). OMIGA uses the CLUSTAL W alignment algorithm using the slow full dynamic programming alignment method with default parameters of an open gap penalty of 10 and an extend gap penalty of 5.0, to find the best alignment between two nucleotide sequences. When using CLUSTAL W or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence such that gaps, mismatches, or insertions of up to 5% of the total number of nucleotides in the reference sequence are allowed.

This embodiment of the present invention is directed to polynucleotides comprising a first nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a nucleic acid sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID

NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or a complementary sequence thereof, irrespective of whether they have functional activity. This is because even where
5 a particular polynucleotide does not have functional activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe, an S1 nuclease mapping probe, or a polymerase chain reaction (PCR) primer.

Preferred, however, are polynucleotides comprising a nucleic acid, the
10 sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a nucleic acid sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18,
15 SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or a complementary sequence thereof, which do, in fact, have functional activity, in particular, polynucleotides which are capable of regulating transcription in *Corynebacterium* species. Examples of polynucleotides which regulate transcription include promoters, operators, transcription factor recognitions sites,
20 transcriptional complexes, and transcriptional regulatory elements, as described herein. Preferably, polynucleotides of the present invention comprise a promoter which functions in *Corynebacterium* species, preferably in *Corynebacterium glutamicum*, *Brevibacterium flavum*, or *Brevibacterium lactofermentum*, even more preferably in *C. glutamicum*. Assays to determine whether a polynucleotide
25 is capable of regulating transcription in *Corynebacterium* species can be routinely performed using techniques described herein and otherwise known in the art.

Accordingly, the present invention also encompasses the above polynucleotide, further comprising a second nucleic acid. Preferably, the second nucleic acid encodes one or more polypeptides, and the physical location of the
30 first nucleic acid relative to the second nucleic acid is such that, under the

appropriate conditions, the first nucleic acid will operably regulate transcription of the second nucleic acid, thereby facilitating production of the one or more polypeptides. "Facilitation of production" includes increasing production constitutively, decreasing or blocking production except under specific conditions, and/or increasing production except under specific conditions. Among the "conditions" contemplated by the present invention are: (1) adding a component to a culture medium, (2) removing a component from a culture medium, (3) replacing one component of a culture medium with a second component, (4) increasing the temperature of the culture medium, (5) decreasing the temperature of the culture medium, and (6) regulating the atmospheric conditions (*e.g.*, oxygen or nitrogen concentrations) in which the culture medium is maintained. Examples of such conditions are described in more detail, *infra*.

One particularly preferred aspect of this embodiment is a polynucleotide comprising a nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a nucleic acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22; and a second nucleic acid, which encodes a component of an amino acid biosynthesis pathway. Examples of amino acid biosynthesis pathways included in this aspect of the invention are the enzymatic pathways that result in the synthesis of L-glycine, L-alanine, L-leucine, L-methionine, L-phenylalanine, L-tryptophan, L-lysine, L-glutamine, L-glutamic acid, L-serine, L-proline, L-valine, L-isoleucine, L-cysteine, L-tyrosine, L-histidine, L-arginine, L-asparagine, L-aspartic acid, and L-threonine.

A preferred amino acid biosynthesis pathway results in the synthesis of L-lysine. In this preferred aspect of this embodiment, the second nucleic acid encodes one of the following components of the L-lysine biosynthesis pathway:

aspartokinase, aspartate beta-semialdehyde dehydrogenase, diaminopimelate dehydrogenase, diaminopimelate decarboxylase, dihydrodipicolinate synthetase, dihydrodipicolinate reductase, or pyruvate carboxylase.

The present invention is further directed to a polynucleotide comprising
5 a first nucleic acid, the sequence of which comprises about 6 contiguous nucleotides, preferably about 10 contiguous nucleotides, even more preferably about 12, 15, 20, 30, 40, 50, 100, 150, 200, or in some cases up to about 500, 1000, or 1500 contiguous nucleotides, of a sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ
10 ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or the complement of any of these sequences. By a first nucleic acid the sequence of which comprises about 10 contiguous nucleotides of any of said
15 nucleic acid sequences, for example, is intended a nucleic acid which includes about 10 contiguous bases from any of said nucleotide sequences. Of course, the polynucleotide includes at least 10 contiguous nucleotides of any of said nucleic acid sequences, and may include about 12, 15, 20, 30, 40, 50, 100, 150, 200, or in some cases up to about 500, 1000, or 1500 contiguous nucleotides, or the entire
20 sequence, of any of said nucleic acid sequences. In this context "about" includes the particularly recited size, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini.

Representative examples of polynucleotides of the invention include a first nucleic acid the sequence of which comprises, for example, a sequence from
25 about nucleotide 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, or 450-500 of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, or SEQ ID NO:22, or the
30 complementary DNA strand thereto; a sequence from about nucleotide 1-50, 51-

100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 450-500,
 501-550, 551-600, 601-650, 651-700, 701-750, 751-800, 801-850, 851-900, 901-
 950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-
 1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500 or 1501-1583 of SEQ ID
 5 NO:9, or the complementary DNA strand thereto; a sequence from about
 nucleotide 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400,
 401-450, 450-500, 501-550, 551-600, 601-650, 651-700, 701-750, 751-800, 801-
 850, 851-900, 901-950, 951-1000, or 1001-1083 of SEQ ID NO:12, or the
 complementary DNA strand thereto; or a sequence from about nucleotide 1-50,
 10 51-100, 101-150, 151-200, 201-250, or 251-297 of SEQ ID NO:18, or the
 complementary DNA strand thereto. In this context "about" includes the
 particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1)
 nucleotides, at either terminus or at both termini. Additional examples of
 polynucleotides of the invention include a first nucleic acid the sequence of which
 15 comprises a consensus -10 or -35 region as listed in Table 2, *infra*.

This embodiment of the present invention is directed to a polynucleotide
 comprising a first nucleic acid, the sequence of which comprises about 6
 contiguous nucleotides, preferably about 10 contiguous nucleotides, even more
 preferably about 12, 15, 20, 30, 40, 50, 100, 150, 200, or in some cases up to
 20 about 500, 1000, or 1500 contiguous nucleotides, of a sequence selected from the
 group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7,
 SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12,
 SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID
 NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and
 25 SEQ ID NO:22, or the complement of any of these sequences, irrespective of
 whether they have functional activity. This is because even where a particular
 polynucleotide does not have functional activity, one of skill in the art would still
 know how to use the nucleic acid molecule, for instance, as a hybridization probe,
 an S1 nuclease mapping probe, or a polymerase chain reaction (PCR) primer.

Preferred, however, a polynucleotide comprising a first nucleic acid, the sequence of which comprises about 6 contiguous nucleotides, preferably about 10 contiguous nucleotides, even more preferably about 12, 15, 20, 30, 40, 50, 100, 150, 200, or in some cases up to about 500, 1000, or 1500 contiguous nucleotides, of a sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or the complement of any of these sequences, which do, in fact, have functional activity, in particular, polynucleotides which regulate transcription in *Corynebacterium* species. Examples of polynucleotides which regulate transcription in *Corynebacterium* species include promoters, operators, transcription factor recognitions sites, transcriptional complexes, and transcriptional regulatory elements, as described herein.

Preferably, polynucleotides of the present invention comprise a promoter which functions in *Corynebacterium* species, preferably in *Corynebacterium glutamicum*, *Brevibacterium flavum*, or *Brevibacterium lactofermentum*, even more preferably in *C. glutamicum*. Assays to determine whether a polynucleotide is capable of regulating transcription in *Corynebacterium* species can be routinely performed using techniques described herein and otherwise known in the art.

Accordingly, the present invention also encompasses the above polynucleotide, further comprising a second nucleic acid. Preferably, the second nucleic acid encodes one or more polypeptides, and the physical location of the first nucleic acid relative to the second nucleic acid is such that, under the appropriate conditions, the first nucleic acid will operably regulate transcription of the second nucleic acid, thereby facilitating production of the one or more polypeptides, as described herein.

One particularly preferred aspect of this embodiment is a polynucleotide comprising a first nucleic acid, the sequence of which comprises about 6 contiguous nucleotides, preferably about 10 contiguous nucleotides, even more preferably about 12, 15, 20, 30, 40, 50, 100, 150, 200, or in some cases up to about 500, 1000, or 1500 contiguous nucleotides, of a nucleic acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22; and a second nucleic acid, which encodes a component of an amino acid biosynthesis pathway. A preferred amino acid biosynthesis pathway results in the synthesis of L-lysine. Examples of amino acid biosynthesis pathways, and specific components of the L-lysine biosynthesis pathway are described herein.

The nucleotide sequences of functional bacterial promoters exhibit a limited but significant degree of variability depending on the gene that is regulated by the particular promoters in question, and the bacterial species in which the promoter is found. See, *e.g.*, Pátek *et al.*, *Microbiology* 142:1297 (1996), which is incorporated herein by reference in its entirety. Despite this variability, investigators have identified promoter consensus sequences by comparing the nucleotide sequences of multiple bacterial promoters to one another. A consensus sequence is a sequence that reflects the most common nucleotides that are found in particular positions for a multitude of promoters. Frequently, the functionality of a promoter consensus sequence is confirmed by mutational analysis in an experimental system; *i.e.*, the expression of a reporter gene that is under the control of a mutated promoter is assessed.

Early experiments and analyses of *E. coli* promoters identified a consensus sequence that comprises two nucleotide hexamers (six contiguous nucleotides) separated by about 17 nucleotides, and located about 35 and 10 nucleotides, respectively, upstream of the transcriptional start site for the

corresponding gene. Pátek *et al.*, *Microbiology* 142:1297 (1996). These two hexamers are known individually as the -35 and -10 regions, and their respective consensus sequences are TTGaca, and TAtaaT. Within a consensus sequence, capital letters denote particular nucleotides that are found in analogous positions in at least 70% of the promoter sequences analyzed. Lower case letters denote nucleotides that are found in analogous positions in between 42% to 70% of the promoter sequences analyzed.

Recently, Pátek *et al.* analyzed the sequences of 18 promoters from *C. glutamicum*, and four promoters from the temperate corynephage ϕ GA1, to identify a promoter consensus sequence specific to *C. glutamicum*. From their analysis, Pátek *et al.* identified a -35 consensus sequence of ttGcca, and a -10 consensus sequence of TA.aaT. The period (".") at the third position of the -10 consensus sequence indicates that the nucleotide at this position does not establish a consensus among the *C. glutamicum* promoters that were compared by Pátek *et al.*

Accordingly, in one aspect of this embodiment of the present invention, a polynucleotide is provided which comprises a -10 promoter consensus sequence, and/or a -35 consensus sequence derived from a sequence selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22. While not being bound by theory, such sequence motifs may function as part of a promoter capable of regulating transcription in *Corynebacterium* species, preferably in *Corynebacterium glutamicum*, *Brevibacterium flavum*, or *Brevibacterium lactofermentum*, even more preferably in *C. glutamicum*.

To identify -10 promoter and/or -35 promoter consensus sequences, computer analyses of SEQ ID NOS: 1 through 22 were conducted using the prokaryotic nucleic acid motif search functionality provided by the OMIGA 2.0

computer software package. In addition, SEQ ID NOs: 1 through 22 were analyzed visually to identify any sequences that agree with the consensus sequences identified by Pátek *et al.* A summary of predicted -10 and -35 consensus sequences identified for SEQ ID NO: 1 through 22 is listed in Table 2.

5 It is to be noted that in the numbering convention of SEQ ID NOs 1 through 22, all nucleotides have a positive number and the numbers increase. This is in contrast to the normal convention for promoter regions, wherein the nucleotides are assigned negative numbers which increase up to nucleotide number 1 at the putative transcription start site.

TABLE 2

SEQ ID NO:	-10 sequence(s)	Position	-35 sequence(s)	Position
1			TTGAGA*	375
1			TTGAAA*	386
3			TTGAAA*	339
4	TACAAT‡	275		
4			TTGAGA*	57
4			TTGAAA*	186
4			TTGAAA*	301
4			TTGCCA‡	466
5			TTGAAA*	307
5			TTGAGA*	340
7	TACAAT‡	334		
7			TTGCCA‡	224
8			TTGATA*	341
8			TTGCCA‡	371
9			TTGATA*	803
10	TACAAT‡	406		
10			TTGACA*	234
11			TTGAAA*	248
13	TACAAT‡	394		
13			TTGCCA‡	118
14	TAAAAT‡	169		
15	TAAAAT‡	379		
17			TTGACA*	211
18			TTGAAA*	31
18			TTGAAA*	65
18	TAAAAT‡	269		
18			TTGCCA‡	168
20			TTGAAA*	490
21			TTGAAA*	41
21			TTGAGA*	445
21			TTGAAA*	494

* consensus sequences identified by the OMIGA 2.0 program

‡ consensus sequences identified visually by comparison to consensus sequences set forth in Pátek *et al.*

In another embodiment, the invention provides an isolated polynucleotide comprising a first nucleic acid at least 10 nucleotides in length, preferably at least 12, 15, 30, 50 or 150 nucleotides in length, which hybridizes under stringent conditions to a reference nucleic acid, or the complement thereof, wherein the sequence of said reference nucleic acid is selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22. By "stringent hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

By a first nucleic acid "at least 10 nucleotides in length," which hybridizes to a reference nucleic acid is intended a first nucleic acid (either DNA or RNA) which hybridizes to 10 or more contiguous nucleotides from the nucleotide sequence of the reference nucleic acid (e.g., SEQ ID NOs 4-22).

This embodiment of the present invention is directed to an isolated polynucleotide comprising a first nucleic acid at least 10 nucleotides in length, preferably at least 12, 15, 30, 50 or 150 nucleotides in length, which hybridizes under stringent conditions to a reference nucleic acid, or the complement thereof, wherein the sequence of said reference nucleic acid is selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or the complement of any of these sequences, irrespective of whether they have functional activity. This is because even where a particular polynucleotide does not have functional activity, one of skill in the art would still know how to

use the nucleic acid molecule, for instance, as a hybridization probe, an S1 nuclease mapping probe, or a polymerase chain reaction (PCR) primer.

Preferred, however, an isolated polynucleotide comprising a first nucleic acid at least 10 nucleotides in length, preferably at least 12, 15, 30, 50 or 150 nucleotides in length, which hybridizes under stringent conditions to a reference nucleic acid, or the complement thereof, wherein the sequence of said reference nucleic acid is selected from the group consisting of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22, or the complement of any of these sequences, which do, in fact, have functional activity, in particular, polynucleotides which regulate transcription in *Corynebacterium* species. Examples of polynucleotides which regulate transcription in *Corynebacterium* species include promoters, operators, transcription factor recognitions sites, transcriptional complexes, and transcriptional regulatory elements, as described herein. Preferably, polynucleotides of the present invention comprise a promoter which functions in *Corynebacterium* species, preferably in *Corynebacterium glutamicum*, *Brevibacterium flavum*, or *Brevibacterium lactofermentum*, even more preferably in *C. glutamicum*. Assays to determine whether a polynucleotide is capable of regulating transcription in *Corynebacterium* species can be routinely performed using techniques described herein and otherwise known in the art.

Accordingly, the present invention also encompasses the above polynucleotide, further comprising a second nucleic acid. Preferably, the second nucleic acid encodes one or more polypeptides, and the physical location of the first nucleic acid relative to the second nucleic acid is such that, under the appropriate conditions, the first nucleic acid will operably regulate transcription of the second nucleic acid, thereby facilitating production of the one or more polypeptides, as described herein.

One particularly preferred aspect of this embodiment is an isolated polynucleotide comprising a first nucleic acid at least 10 nucleotides in length, preferably at least 12, 15, 30, 50 or 150 nucleotides in length, which hybridizes under stringent conditions to a reference nucleic acid, or the complement thereof, wherein the sequence of said reference nucleic acid is selected from the group consisting of selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, and SEQ ID NO:22; and a second nucleic acid, which encodes a component of an amino acid biosynthesis pathway. A preferred amino acid biosynthesis pathway results in the synthesis of L-lysine. Examples of amino acid biosynthesis pathways, and specific components of the L-lysine biosynthesis pathway are disclosed herein.

Another embodiment of the invention provides an isolated polynucleotide comprising a first nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 2, and a second nucleic acid which encodes a polypeptide which functions as a component of the lysine biosynthesis pathway. Alternatively in this embodiment, the sequence of the first nucleic acid comprises about 6 contiguous nucleotides, preferably about 10 contiguous nucleotides, even more preferably about 12, 15, 20, 30, 40, 50, 100, 150, 200, or 500, contiguous nucleotides, of SEQ ID NO:2, or the first nucleic acid is at least 10 nucleotides in length, preferably at least 12, 15, 30, 50 or 150 nucleotides in length, and hybridizes under stringent conditions to a reference nucleic acid, or the complement thereof, wherein the sequence of said reference nucleic acid is SEQ ID NO:2. A preferred amino acid biosynthesis pathway results in the synthesis of L-lysine. Examples of components of the L-lysine biosynthesis pathway are disclosed herein. In this embodiment, the first nucleic acid, which comprises the *C. glutamicum aceA* promoter, regulates transcription of the second nucleic acid.

Yet another embodiment of the invention provides an isolated polynucleotide comprising a first nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 3; and a second nucleic acid which encodes polypeptide which functions as a component of an amino acid biosynthesis pathway. Alternatively in this embodiment, the sequence of the first nucleic acid comprises about 6 contiguous nucleotides, preferably about 10 contiguous nucleotides, even more preferably about 12, 15, 20, 30, 40, 50, 100, 150, 200, or 500, contiguous nucleotides, of SEQ ID NO:3, or the first nucleic acid is at least 10 nucleotides in length, preferably at least 12, 15, 30, 50 or 150 nucleotides in length, and hybridizes under stringent conditions to a reference nucleic acid, or the complement thereof, wherein the sequence of said reference nucleic acid is SEQ ID NO:3. A preferred amino acid biosynthesis pathway results in the synthesis of L-lysine. Examples of amino acid biosynthesis pathways, and specific components of the L-lysine biosynthesis pathway are disclosed herein. In this embodiment, the first nucleic acid, which comprises the *C. glutamicum aceB* promoter, regulates transcription of the second nucleic acid.

A further embodiment of the invention provides an isolated *Corynebacterium* chromosome with a first nucleic acid integrated into the chromosome, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:1. By "an isolated *Corynebacterium* species chromosome" is meant a bacterial chromosome which has been altered in some way to affect gene expression, either of naturally-occurring genes or heterologous genes, and then isolated from other non-altered chromosomes, *e.g.*, by selection and isolation of a host cell harboring the chromosome. In this embodiment, the isolated *Corynebacterium* species chromosome also has a second nucleic acid integrated therein, wherein said second nucleic acid encodes polypeptide which functions as a component of an amino acid biosynthesis pathway. A preferred amino acid biosynthesis pathway results in the synthesis of L-lysine. Examples of amino acid biosynthesis

pathways, and specific components of the L-lysine biosynthesis pathway are disclosed herein. In this embodiment, the first nucleic acid, which comprises the *C. glutamicum pta* promoter, regulates transcription of the second nucleic acid.

Another embodiment of the invention provides an isolated
5 *Corynebacterium* chromosome with a first nucleic acid integrated into the chromosome, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:23. In this embodiment, the isolated *Corynebacterium* species chromosome also has a second nucleic acid integrated therein, wherein said second nucleic acid encodes polypeptide which
10 functions as a component of an amino acid biosynthesis pathway, preferably an L-lysine biosynthesis pathway. In this embodiment, the first nucleic acid, which comprises the *E. coli araBAD* promoter and the *araC* regulator molecule, regulates transcription of the second nucleic acid. The *araC/araBAD* transcriptional complex is a tightly regulated system responsive to the regulator
15 molecule arabinose.

Yet another embodiment of the invention provides an isolated *Corynebacterium* chromosome with a first nucleic acid integrated into the chromosome, the sequence of which is identical to either SEQ ID NO:26 or SEQ ID NO:27. As with the embodiments described above, the isolated
20 *Corynebacterium* species chromosome also has a second nucleic acid integrated therein, wherein said second nucleic acid encodes polypeptide which functions as a component of an amino acid biosynthesis pathway, preferably an L-lysine biosynthesis pathway, and the first nucleic acid, which is a *trc* or *tac* synthetic promoter, regulates transcription of the second nucleic acid. These synthetic
25 promoters combine the -35 region from the *trp* promoter with the -10 region from the *lacUV5* promoter. See Brosius, J. et al., *J. Biol. Chem.* 260, 3539 (1985). In the embodiment, the isolated chromosome further has a third nucleic acid integrated therein, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:28. This third nucleic
30 acid is operably linked to a transcription control region, and the nucleic acid

encodes a polypeptide at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO:29, the *E. coli lac* repressor gene. The *lacI^q-trc* (SEQ ID NO:30) or the *lacI^q-tac* transcriptional complex is responsive to the regulatory molecules isopropylthiogalactoside (IPTG) and lactose. This regulatory complex in *Corynebacterium glutamicum*, however, is leaky, *i.e.*, transcription occurs in the absence of any added inducer.

In each of the isolated chromosomes described above, the second nucleic acid may be a naturally occurring *Corynebacterium* nucleic acid, situated in its normal position on the chromosome, it may be a naturally occurring *Corynebacterium* nucleic acid which has been moved to a different position on the chromosome, or it may be a non-native nucleic acid. The first nucleic acid is engineered to be in a physical location which allows it to regulate transcription of the second nucleic acid either through *in vitro* cloning of a vector having both the first and second nucleic acids followed by homologous or random recombination into the *Corynebacterium* species chromosome, or it is inserted in front of a selected second nucleic acid already situated in the chromosome through homologous recombination. Such manipulations are readily understood by those of ordinary skill in the art.

In another aspect of the above embodiments, a *Corynebacterium* species host cell comprising a *Corynebacterium* species chromosome as described above is provided, preferably a host cell derived from *Corynebacterium glutamicum*, *Brevibacterium flavum*, and *Brevibacterium lactofermentum*, and even more preferably derived from *Corynebacterium glutamicum*. In a related aspect, a method of producing such a *Corynebacterium* species host cell is provided, which comprises transforming *Corynebacterium* species cells with a vector comprising a first nucleic acid as described above, wherein said vector facilitates integration of the first nucleic acid into the chromosome of said *Corynebacterium* species cells, and selecting the host cell in certain preferred embodiments, the vector also comprises the second nucleic acid described above, and/or the third nucleic acid

described above, with the first nucleic acid being physically situated to regulate transcription of the second nucleic acid. These vectors are also provided.

The present invention further relates to variants of the polynucleotides of the present invention. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. *Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions which may involve one or more nucleotides. Particularly preferred are variants which alter the level of activity of the polynucleotide relative to a non-variant polynucleotide. For example, variants which increase or decrease the strength of a promoter, or variants which alter or eliminate the need for an inducer. Methods to construct variant promoters and screen them for their relative strengths utilizing reporter genes are well known in the art. *See, e.g.,* Vašicová, *et al.*, *J. Bacteriol.* 181:6188-6199 (1999), which is incorporated herein by reference in its entirety.

D. Vectors and Host Cells

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are engineered with vectors of the invention and methods to increase the production of metabolic products such as amino acids purines, pyrimidines, and heterologous polypeptides through utilization of host cells of the invention.

Host cells can be engineered to incorporate polynucleotides of the present invention. The polynucleotides may be introduced alone or with other polynucleotides. Such other polynucleotides may be introduced independently, co-introduced or introduced joined to the polynucleotides of the invention.

In accordance with this aspect of the invention the vector may be, for example, a plasmid vector or a bacteriophage vector. Such vectors may be introduced into *Corynebacterium* cells such that they are maintained and replication as an extrachromosomal element, or, alternatively, such that they
5 integrate into the chromosome. Vectors are introduced into host cells by well known techniques for introducing DNA and RNA into prokaryotic cells, examples of which are described in the Examples, *infra*.

Preferred vectors comprise a first nucleic acid which is capable of regulating transcription of a second nucleic acid which is inserted in operable
10 association, preferably immediately downstream, of the first nucleic acid. The second nucleic acid preferably encodes a polypeptide such as an enzyme in an amino acid biosynthesis pathway. Especially preferred among vectors are those which allow convenient insertion of the second nucleic acid into the vector into the proper position to be regulated by the first nucleic acid, for example, through
15 the use of a multiple cloning region. Appropriate trans-acting factors such as inducers either are supplied by the host cell, supplied by a complementing vector or supplied by the vector itself upon introduction into the host cell.

A great variety of vectors can be used in the invention. Such vectors include chromosomal, episomal and virus-derived vectors *e.g.*, vectors derived
20 from bacterial plasmids and from bacteriophage, as well as vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used in accordance with this aspect of the present invention. Generally, any vector suitable to maintain and propagate a polynucleotide in a bacterial host may be used in this
25 regard.

A large numbers of suitable vectors for use in bacteria are known, many of which are commercially available. Preferred prokaryotic vectors include plasmids such as those capable of replication in *E. coli* (such as, for example, pBR322, ColEI, pSC101, pACYC 184, π VX. Such plasmids are, for example,
30 disclosed by Maniatis, T., *et al.*, *In: Molecular Cloning*, A Laboratory Manual,

Cold Spring Harbor Press, Cold Spring Harbor, NY (1982)). The following vectors are provided by way of example: pET (Novagen), pQE70, pQE60, pQE-9 (Qiagen), pBs, phagescript, psiXI74, pBlueScript SK, pBsKS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia).

Preferred vectors for the isolated polynucleotides of the invention include the pFC1 to pFC7 novel family of combinatorial cloning vectors (Lonsdale, D.M., *et al.*, *Plant Molecular Biology Reporter* 13: 343-345 (1995)), the pK184 vector (Jobling, M.G. and Homes, R.K., *Nucleic Acid Research* 18: 5315-5316 (1990)).

Another group of preferred vectors are those that are capable of autonomous replication in *Corynebacterium* species. Such vectors are well known to those skilled in the art of amino acid production by way of microbial fermentation, examples of which include pSR1, pMF1014a and vectors derived therefrom.

The second nucleic acid encoding a polypeptide is operatively associated with the first nucleic acid as described above, which provides the appropriate transcription control sequence(s)), in particular, a promoter. In addition, the vector may contain sites for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The second nucleic acid, which provides the coding region of the polypeptide to be regulated will include a translation initiating AUG or GUG at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the coding region.

In addition, and as described in more detail herein, the vector may contain control regions that regulate as well as engender expression. Generally, such regions will operate by controlling transcription, such as inducer or repressor binding sites and enhancers, among others.

Vectors of the present invention generally will include a selectable marker. Such markers also may be suitable for amplification or the vectors may contain additional markers for this purpose. In this regard, vectors preferably

contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Such markers include, but are not limited to, an antibiotic resistance gene such as a chloramphenicol, ampicillin, or kanamycin resistance gene, or an autotrophic gene which allows the host cell to grow in the absence of a nutrient for which the host cell strain is normally auxotrophic.

If the vector is intended to be maintained in the host cell extrachromosomally, it will contain, in addition and origin of replication which will allow it to replicate in the *Corynebacterium* species host cell. Alternatively, if it is desired that the vector integrate into the *Corynebacterium* species chromosome, the vector is constructed such that it cannot replicate in *Corynebacterium*. For example, such a vector might be capable of propagation in another organism, for example, *E. coli*, but lack the proper origin of replication to be propagated in *Corynebacterium*. In another aspect of this embodiment, the vector is a shuttle vector which can replicate and be maintained in more than one host cell species, for example, such a shuttle vector might be capable of replication in a *Corynebacterium* host cell such as a *C. glutamicum* host cell, and also in an *E. coli* host cell.

A vector comprising a polynucleotide as described elsewhere herein may be introduced into an appropriate *Corynebacterium* species host cell using a variety of well known techniques, for example, transformation or electroporation. Appropriate culture mediums and conditions for culturing *Corynebacterium* host cells are known in the art.

The invention provides methods for increasing amino acid production and processes for the production of an amino acid wherein increased biosynthetic pathway gene expression is accomplished through insertion of an isolated polynucleotide into the chromosome of a host cell. For example, insertion of isolated polynucleotides into the chromosome of *Corynebacterium* species may be done utilizing the pK184 plasmid described by Jobling, M. *et al.*, *Nucleic Acids Research* 18(17): 5315-5316 (1990). Because these vectors lack a *Corynebacterium* species origin of replication and contains a selectable marker

such as kanamycin (*kan*), cells will only be capable of growing under selection if the vector has been inserted into the host cell chromosome by homologous recombination.

5 The invention also provides methods for increasing amino acid production and processes for the production of an amino acid wherein increased biosynthetic pathway gene expression is accomplished through the introduction into a host cell of a self-replicating, extra-chromosomal vector, *e.g.*, a plasmid, comprising an isolated nucleic acid molecule encoding an amino acid biosynthetic pathway gene or genes. Suitable plasmids for these embodiments include pSR1 and other
10 derivatives of pSR1 (Archer, J. *et al.*, *J. Gen. Microbiol.* 139: 1753-1759 (1993)).

These vectors are listed solely by way of illustration of the many vectors available to those of skill in the art. Other vectors are described in the Examples, *infra*. Selection of appropriate vectors to transform into a host cell is a well known procedure and the requisite techniques for vector construction,
15 introduction of the vector into the host and maintenance in the host are routine skills in the art.

The present invention also relates to *Corynebacterium* species host cells comprising the above-described vector constructs described herein. The host cell can be any *Corynebacterium* species, preferably *Corynebacterium glutamicum*,
20 *Brevibacterium flavum*, or *Brevibacterium lactofermentum*, even more preferably a strain of *C. glutamicum*. Preferred strains of *C. glutamicum* to use as host cells of the invention include: NRRL-B11474, ATCC 21799, ATCC 21529, ATCC 21543, and E12.

25 Introduction of a vector into the host cell can be effected by transformation, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis *et al.*, *Basic Methods in Molecular Biology* (1986).

E. Methods and Processes of the Invention

Various embodiments of the invention provide methods to increase the production of an amino acid and processes for the production of an amino acid from a *Corynebacterium* species host cell. Particularly preferred *Corynebacterium* species of the methods and processes of the invention include: *Corynebacterium glutamicum*, *Brevibacterium flavum*, *Brevibacterium lactofermentum* and other *Cornynebacteria* and *Brevibacteria* species known in the art.

Amino acid biosynthetic pathway genes embodied by the methods and processes described herein include those for L-glycine, L-alanine, L-methionine, L-phenylalanine, L-tryptophan, L-proline, L-serine, L-threonine, L-cysteine, L-tyrosine, L-asparagine, L-glutamine, L-aspartic acid, L-glutamic acid, L-lysine, L-arginine, L-histidine, L-isoleucine, L-leucine, and L-valine biosynthesis. Particularly preferred embodiments are drawn to biosynthetic pathway genes for L-lysine (Sahm *et al.*, *Ann. N. Y. Acad. Sci.* 782: 25-39 (1996)), L-threonine, L-isoleucine, L-tryptophan, and L-valine.

By way of example, the amino acid pathway for L-lysine biosynthesis is well known to skilled artisans of amino acid production in *Corynebacterium* species. Genes encoding the enzymes important for the conversion of L-aspartate to L-lysine include the *ask*, *asd*, *dapA*, *dapB*, *ddh*, and *lysA* genes, which encode, respectively, aspartokinase, aspartate beta-semialdehyde dehydrogenase, dihydrodipicolinate synthetase, dihydrodipicolinate reductase, diaminopimelate dehydrogenase, diaminopimelate decarboxylase, and pyruvate carboxylase. Thus, the invention provides herein for exemplary purposes only, specific embodiments utilizing L-lysine biosynthetic pathway genes. Other embodiments drawn to the use of biosynthetic pathway genes for the synthesis of other amino acids are also encompassed by the invention described herein.

The methods to increase the production of a biosynthetic product such as an amino acid and the processes for the production of the biosynthetic product of

the invention involve increasing the expression of at least one biosynthetic pathway gene. Preferred methods of increasing expression comprise using promoters of the invention.

Accordingly, the present invention provides a first nucleic acid which preferably comprises a promoter. Preferred promoters include, but are not limited to the following types of promoters: a constitutive promoter, an inducible promoter, a derepressable promoter, a heat sensitive promoter, and a cold sensitive promoter. In this aspect of the invention, a constitutive promoter causes the second nucleic acid to be transcribed at a substantially constant rate; an inducible promoter will be non-functional or will function at a reduced level under certain standard culture conditions, but causes elevated transcription of the second nucleic acid when the transformed host cells are grown in the presence of an appropriate inducer substance; a derepressable promoter causes elevated transcription of the second nucleic acid when the transformed host cells are grown in the absence of an appropriate suppressor substance, but will be non-functional or will function at a reduced level when the suppressor substance is present; a heat sensitive promoter causes elevated transcription of the second nucleic acid when the transformed host cells are grown in medium the temperature of which is adjusted to a temperature that is greater than the optimum temperature for wild-type untransformed host cells; and a cold sensitive promoter causes elevated transcription of the second nucleic acid when said transformed host cells are grown in a medium the temperature of which is adjusted to a temperature that is less than the optimum temperature for a wild-type untransformed host cell.

Also included within this embodiment of the invention are methods wherein the transformed host cells are grown to a sufficient cell density according to methods well known in the art. Further included in this embodiment are methods for substantially purifying the biosynthetic product from said culture, wherein said purification methods are also well known in the art.

In certain embodiments of the present invention, an inducible promoter is used. When an inducible promoter is used, it is preferred that the transformed *Corynebacterium* species host cells are first allowed to proliferate to a sufficient cell density in the absence of the inducer, followed by addition of the inducer to the culture medium thereby stimulating production of said product. Examples of inducers include, but are not limited to one or more of the following substances: acetic acid, pyruvate, ethanol, fatty acids, cellulose subunits, starch subunits, triglycerides, or any of the following sugars: fructose, maltose, lactose, or arabinose.

In other embodiments, a derepressable promoter is used. When a derepressable promoter is used, it is preferred that the transformed *Corynebacterium* species host cells are first allowed to proliferate to a sufficient cell density in the presence of the suppressor, followed by either removal of the suppressor from the culture, or replacement of the medium containing the suppressor with medium lacking the suppressor, thereby stimulating production of said product. Under certain conditions, the suppressor is removed from the medium by being metabolized by the host cell. Examples of said suppressors include, but are not limited to any of the following substances: a purine, a pyrimidine, oxygen, or a suppressor amino acid including histidine.

In other embodiments, the promoter is self-inducing, *i.e.*, the inducer molecule is produced by the host cell and builds up through the course of the growth, thereby changing gene expression as growth progresses. In this embodiment, the promoter is induced either gradually or at some point in the growth of the host cell culture, however this induction does not require the addition of exogenous compounds to the culture medium.

In yet another embodiment, a heat sensitive promoter is used. When a heat sensitive promoter is used, it is preferred that the transformed host cells are first allowed to proliferate to a sufficient cell density in culture medium, the temperature of which is adjusted to the optimum growth temperature for untransformed host cells, followed by adjusting the temperature of said culture

medium to a temperature that is greater than the optimum growth temperature for untransformed host cells.

In yet another embodiment, a cold sensitive promoter is used. When a cold sensitive promoter is used, it is preferred that the transformed host cells are first allowed to proliferate to a sufficient cell density in culture medium, the temperature of which is adjusted to the optimum growth temperature for untransformed host cells, followed by adjusting the temperature of said culture medium to a temperature that is less than the optimum growth temperature for untransformed host cells.

In a related embodiment of the invention, a method is provided for the production of a product such as an amino acid, a purine nucleotide, or a heterologous polypeptide in a *Corynebacterium* species host cell utilizing a heterologous transcriptional complex. According to this embodiment, the host cell harbors the heterologous transcriptional complex integrated into the bacterial chromosome. Preferred heterologous transcriptional complexes include the *E. coli araC-araBAD* transcriptional complex (SEQ ID NO:23), and the *lacI^q-trc* or *lacI^q-tac* transcriptional complexes which combine the *E. coli lacI^q* repressor gene (*lacI^q*, SEQ ID NO:28) with either the *trc* promoter (SEQ ID NO:26 or SEQ ID NO:30) or the *tac* promoter (SEQ ID NO:27). These synthetic promoters combine the -35 region from the *trp* promoter with the -10 region from the *lacUV5* promoter. See Brosius, J. et al., *J. Biol. Chem.* 260, 3539 (1985).

The method utilizing the *E. coli araC-araBAD* transcriptional complex preferably further comprises culturing the host cell in or on a medium to which the inducer arabinose has been added, either throughout growth of the host cells, or at a point during growth where the host cell culture has reached an optimal density. The method utilizing either the *lacI^q-trc* or the *lacI^q-tac* transcriptional complex may further comprise culturing the host cell in or on a medium to which the inducer isopropylthiogalactoside (IPTG) or the inducer lactose has been added, either throughout growth of the host cells, or at a point during growth where the host cell culture has reached an optimal density. However, in *C.*

glutamicum, the *lacI^q/trc* system is leaky, therefore an inducer is not required for optimal growth.

A preferred aspect of this embodiment provides a method for the production of an amino acid in a *Corynebacterium* host cell utilizing a heterologous transcriptional complex integrated into the host cell chromosome, where the heterologous transcriptional complex operably regulates production a polypeptide which functions in an amino acid biosynthesis pathway. In this aspect, the heterologous transcriptional complex, when integrated into the host cell chromosome, regulates transcription of one or more metabolic pathway enzymes, thereby resulting in elevated production of the amino acid by the transformed host cell.

One aspect of this embodiment involves addition of an inducer to the culture medium thereby upregulating the heterologous transcriptional complex. In the method utilizing the *E. coli araC-araBAD* transcriptional complex arabinose is added to the medium. In the method utilizing the *lacI^q-trc* or the *lacI^q-tac* transcriptional complex, IPTG and/or lactose is added to the culture medium. Preferably in this method, the host cells are grown to a sufficient cell density according to methods well known in the art prior to addition of the inducer molecule.

For those embodiments of the invention drawn to a method to increase production of an amino acid, screening for increased production of an amino acid, for example L-lysine, may be determined by directly comparing the amount of L-lysine produced in culture by a standard *Corynebacterium* species strain to that of a *Corynebacterium* species host cell which has been transformed to comprise an isolated polynucleotide of the present invention. The level of production of the amino acid of choice may conveniently determined by the following formula to calculate the percent yield from dextrose: $[(\text{g amino acid/L} / (\text{g dextrose consumed/L})) * 100]$.

A variety of media known to those skilled in the art may be used to support cell growth for the production of an amino acid. Illustrative examples of

suitable carbon sources include, but are not limited to: carbohydrates, such as glucose, fructose, sucrose, starch hydrolysate, cellulose hydrolysate and molasses; organic acids, such as acetic acid, propionic acid, formic acid, malic acid, citric acid, and fumaric acid; and alcohols, such as glycerol. Illustrative examples of
5 suitable nitrogen sources include, but are not limited to: ammonia, including ammonia gas and aqueous ammonia; ammonium salts of inorganic or organic acids, such as ammonium chloride, ammonium phosphate, ammonium sulfate and ammonium acetate; and other nitrogen-containing sources, including meat extract, peptone, corn steep liquor, casein hydrolysate, soybean cake hydrolysate,
10 urea and yeast extract.

A variety of fermentation techniques are known in the art which may be employed in processes of the invention drawn to the production of amino acids. Generally, amino acids may be commercially produced from the invention in fermentation processes such as the batch type or of the fed-batch type. In batch
15 type fermentations, all nutrients are added at the beginning of the fermentation. In fed-batch or extended fed-batch type fermentations one or a number of nutrients are continuously supplied to the culture, right from the beginning of the fermentation or after the culture has reached a certain age, or when the nutrient(s) which are fed were exhausted from the culture fluid. A variant of the extended
20 batch of fed-batch type fermentation is the repeated fed-batch or fill-and-draw fermentation, where part of the contents of the fermenter is removed at some time, for instance when the fermenter is full, while feeding of a nutrient is continued. In this way a fermentation can be extended for a longer time.

Another type of fermentation, the continuous fermentation or chemostat
25 culture, uses continuous feeding of a complete medium, while culture fluid is continuously or semi-continuously withdrawn in such a way that the volume of the broth in the fermenter remains approximately constant. A continuous fermentation can in principle be maintained for an infinite time.

In a batch fermentation an organism grows until one of the essential
30 nutrients in the medium becomes exhausted, or until fermentation conditions

-57-

become unfavorable (e.g. the pH decreases to a value inhibitory for microbial growth). In fed-batch fermentations measures are normally taken to maintain favorable growth conditions, e.g. by using pH control, and exhaustion of one or more essential nutrients is prevented by feeding these nutrient(s) to the culture.

5 The microorganism will continue to grow, at a growth rate dictated by the rate of nutrient feed. Generally a single nutrient, very often the carbon source, will become limiting for growth. The same principle applies for a continuous fermentation, usually one nutrient in the medium feed is limiting, all other nutrients are in excess. The limiting nutrient will be present in the culture fluid
10 at a very low concentration, often unmeasurably low. Different types of nutrient limitation can be employed. Carbon source limitation is most often used. Other examples are limitation by the nitrogen source, limitation by oxygen, limitation by a specific nutrient such as a vitamin or an amino acid (in case the microorganism is auxotrophic for such a compound), limitation by sulphur and
15 limitation by phosphorous.

The amino acid may be recovered by any method known in the art. Exemplary procedures are provided in the following: Van Walsem, H.J. & Thompson, M.C., *J. Biotechnol.* 59:127-132 (1997), and U.S. Pat. No. 3,565,951, both of which are incorporated herein by reference.

20 *Examples*

Example 1

Production of Host Cells Utilizing Transcriptional Regulatory Regions of the Present Invention

25 *Corynebacterium* species host cells, in which to express metabolic polypeptides are constructed as follows. Polynucleotides of the present invention which comprise transcriptional regulatory regions related to SEQ ID NOs 1-23, 26-28, and 30 are inserted upstream of a multiple cloning region, thereby

allowing convenient insertion coding regions to be regulated and expressed. Testing of the various transcriptional regulatory regions is conveniently carried by inserting the various regulatory regions in operable association to a known reporter gene, such as β -galactosidase (*lacZ*). Methods and techniques common to the art of recombinant DNA technology are used in making vectors of the invention, as may be found in the many laboratory manuals cited and incorporated herein, for example as found in J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). Construction of exemplary vectors is set forth in Example 2, *infra*.

For cell transformation experiments with the isolated nucleic acid molecules of the invention, the growth and preparation of competent cells may be done according to the following procedure: (1) picking a fresh, single colony of *Corynebacterium glutamicum* and growing a culture overnight in 10 mL CM (SM1) in a 250 mL shake flask at 30 degrees Celsius with agitation; (2) inoculating 200 mL of "Growth Media" with the overnight culture to an optical density (O.D.) of 660 nm of 0.1 in a 500 mL shake flask; (3) growing the culture at 30 degrees Celsius with agitation for 5-6 hours; (4) pouring the culture into a chilled, sealed, sterile 250 mL centrifuge bottle; Spin at 8-10K for ten minutes in Refrigerated Sorvall at 4 degrees Celsius; (5) pouring off the supernatant thoroughly and resuspending the cell pellet in an equal volume of ice-cold, sterile, deionized water; (6) centrifuging the sample again under the same conditions; (7) repeating the water wash remembering to keep everything ice-cold; (8) pouring off the supernatant thoroughly and resuspending the cell pellet in 1 mL of ice-cold, sterile 10% glycerol and transferring the cells to a chilled, sterile, 1.5 mL microcentrifuge tube; (9) spin the sample for 10 minutes in a refrigerated centrifuge; (10) pipetting off and discarding the supernatant, and resuspending the pellet in two to three times the pellet volume (200-400 μ L) of 10% glycerol; and (11) aliquoting, if necessary, the cells into chilled tubes and freezing at -70 Celsius.

Plasmid DNAs are introduced into *Corynebacterium glutamicum* host cells by the following electroporation procedure: (1) pipetting 35 μ L cell/glycerol solution onto the side wall of a chilled 0.1cm electrocuvette; (2) pipetting about 2-4 μ L of plasmid into the solution and mixing the sample by gentle pipetting up and down; (3) bringing the entire solution to the bottom of the electrocuvette by gentle tapping, avoiding the creation of bubbles; (4) keeping the sample on ice until ready for the electroschock step, wiping off any moisture on the outside of the electrocuvette prior to the electroschock administration, and shocking the cells one time at 1.5kV, 200 Ω , 25 μ F.

Cells are allowed to recover from electroporation by: (1) immediately pipetting 1 mL of warm "Recovery Media" into the electrocuvette and thoroughly mixing the solution by pipetting; (2) incubating the solution (in the electrocuvette) at 30 degrees Celsius for at least three hours for antibiotic resistance expression and cell recovery and (3) plating on selection media and incubating at 30 degrees Celsius for 3 days.

Example 2

Preparation of L-Lysine Pathway Constructs Utilizing the Regulatory Regions of the Present Invention

Isolated nucleic acid molecules encoding L-lysine amino acid biosynthesis pathway genes are isolated by methods known to those of ordinary skill in the art.

Constructs which facilitate expression and regulation of L-lysine amino acid biosynthesis pathway genes are introduced into *Corynebacterium* strains by methods such as those described in Example 1. Any strain of *Corynebacterium*, particularly that of *Corynebacterium glutamicum*, may be utilized for the isolation of nucleic acid molecules for use in the preparation of vectors to improve amino acid biosynthetic pathway gene expression. Particularly preferred strains include: NRRL-B11474, ATCC 21799, ATCC 21529, ATCC 21543, and E12. As one skilled in the art would know, the invention is not limited to these specific strain origins.

-60-

Methods and techniques common to the art of recombinant DNA technology are used in making vectors of the invention, as may be found in the many laboratory manuals cited and incorporated herein.

The polymerase chain reaction (PCR) technique may be used in the making of vectors of the invention. In a typical reaction, the standard 10X stock solution (100 mM Tris-HCL, pH 8.3, 500 mM KCL, 1.5 mM MgCl₂) is diluted to 1X for use. Typical reaction conditions were used for PCR amplification: 10 mM Tris, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatin, 200uM deoxynucleotides, 0.2-1.0uM primers and 2.0 U/100ul pfu polymerase. Standard cycling parameters were also employed in PCR reactions: For 30 cycles, template denaturation was performed at 94°C for 30 sec.; 65°C annealing temperature was performed for 1 min (or annealing temperature appropriate for particular primer pair); product extension was performed at 72°C for 1 min (if product is <500 bp), 3 min (if product is >500 bp); and at the end of cycling, a final extension at 72°C for 7 min was performed.

(a) Construction of pTrcLacZ

A plasmid construct designed to replicate as a plasmid in *Corynebacterium*, which contains the β -galactosidase (*lacZ*) coding region operably associated with the *lacI*^q/*trc* regulatory region, was constructed as follows. Plasmid pZap2, a plasmid which is capable of replicating in *Corynebacterium*, was digested with *Xba*I and *Sma*I, which linearize the plasmid just upstream of the *lacZ* gene. A cassette containing *lacI*^q and the *trc* promoter was excised from plasmid pTrc99A, available from Amersham Pharmacia Biotech, by digestion with *Sph*I, filling in the overhang, and then digesting with *Xba*I. The 1517-bp cassette (SEQ ID NO:30) was isolated from an agarose gel, and ligated into the linearized pZap2 plasmid. The resulting ligation mix was used to transform *Escherichia coli*, and colonies were isolated from LB plates supplemented with ampicillin (100 μ g/ml), Xgal, and IPTG (LBXIA medium).

Blue colonies were selected and plasmid isolates were verified by restriction digestion and electrophoresis.

(b) Construction of pAraBADLacZ

A plasmid construct designed to replicate as a plasmid in *Corynebacterium*, which contains the *lacZ* coding region operably associated with the *araC/araBAD* regulatory region, was constructed as follows. Plasmid pZap2 was digested with *Xba*I and *Sma*I as in (a). A cassette containing the *araC* gene and the pBAD promoter was excised from plasmid pBAD18-Cm, available from the American Type Culture Collection, Catalog No. 87396, by digestion with *Cla*I followed by a fill-in reaction and then digestion with *Xba*I. The cassette (*i.e.*, nucleotides 1 to 1338 of SEQ ID NO:23) was isolated from an agarose gel and was ligated into the linearized pZap2 plasmid. Transformed *E. coli* colonies were plated on LBXIA medium as in (a), blue colonies were selected, and plasmid isolates were verified by restriction digestion and electrophoresis.

(c) Construction of pAceB10Zap2

A plasmid construct designed to replicate as a plasmid in *Corynebacterium*, which contains the *lacZ* coding region operably associated with the *C. glutamicum aceB* regulatory region, was constructed as follows. A cassette containing the *aceB* promoter was PCR amplified from plasmid pSLD5 using forward primer 5' GTGCGGATCC TGGCTTTCCA ACGTTT 3' (SEQ ID NO:31, *Bam*HI site underlined) and reverse primer 5' CATGGG CTCGAGATGA CCTGTGCCTA 3' (SEQ ID NO:32, *Sac*I site underlined). The resulting PCR product was digested with *Bam*HI and *Sac*I and ligated into the polylinker region of cloning vector pD11 to produce pD11*aceB*. The resulting plasmid was screened for orientation, and then the *aceB* promoter cassette was released from the vector by digestion with *Kpn*I and *Xba*I and cloned into the *Kpn*I and *Xba*I

sites of plasmid pZap2 as described in (a). Transformed *E. coli* colonies were plated on LBXIA medium as in (a), blue colonies were selected, and plasmid isolates were verified by restriction digestion and electrophoresis.

(d) Construction of 131aceB/ask

5 A plasmid construct designed to integrate into the *C. glutamicum* chromosome, which contains the *C. glutamicum* aspartokinase (*ask*) coding region operably associated with the *C. glutamicum* *aceB* regulatory region, was constructed as follows. Plasmid pBGS131 (described in Spratt, B.G. *et al. Gene* 41:337-342 (1986)) was digested with *Sma*I. The *aceB* cassette was released
10 from plasmid pD11*aceB*, produced as in (c), by digestion with appropriate restriction enzymes and the cassette was subjected to a fill-in reaction to produce blunt ends. The cassette was then ligated into the digested pBGS131. Plasmid isolates were tested for the proper orientation of the *aceB* promoter by restriction digestion and electrophoresis, resulting in plasmid 131*aceB*. The *C. glutamicum*
15 *ask* coding region was released from pTrc99A*ask* (produced as in (f), *infra*) by digestion with *Sac*I and *Kpn*I and cloned into the *Sac*I and *Kpn*I sites of 131*aceB*, to produce 131*aceBask*.

(e) Construction of PD10aceBask

20 A plasmid construct designed to replicate as a plasmid in *C. glutamicum*, which contains the *C. glutamicum* aspartokinase (*ask*) coding region operably associated with the *C. glutamicum* *aceB* regulatory region, was constructed as follows. A cassette containing the *ask* coding region operably associated with the *aceB* regulatory region was released from 131*aceBask*, produced as in (d), by digestion with *Bam*HI and *Kpn*I. The resulting fragment was cloned into the
25 *Bam*HI and *Kpn*I sites of plasmid pD10, which is capable of replicating in *C. glutamicum*, to produce pD10*aceBask*.

(f) *Construction of 131LacI^q/trc-ask*

A plasmid construct designed to integrate into the *C. glutamicum* chromosome, which contains the *C. glutamicum* aspartokinase (*ask*) coding region operably associated with the *lacI^q/trc* regulatory region, was constructed as follows. The PCR-amplified *ask* coding region was digested with *SacI* and *KpnI*, and ligated into plasmid pTrc99A which had been digested with *SacI* and *KpnI* to produce pTrc99A*ask*. A cassette containing the *ask* coding region operably associated with the *lacI^q/trc* regulatory region was then released from pTrc99A*ask* by digestion with *SphI*, and was cloned into the *SphI* site of plasmid pBGS131 to produce 131LacI^q/trc-*ask*.

(g) *Construction of Additional LacZ Plasmid Expression Constructs*

Plasmid constructs designed to replicate as a plasmid in *Corynebacterium*, which contain the *lacZ* coding region operably associated with *C. glutamicum* regulatory regions contained in SEQ ID NOs 1, 2, and 4-22, are constructed as follows. Cassettes containing the regulatory regions present in SEQ ID NOs 1, 2, and 4-22 are PCR amplified from the *C. glutamicum* chromosome using primers engineered to introduce appropriate restriction enzyme recognition sites onto the ends of the cassettes. In designing the primers, care is taken to introduce sites for restriction enzymes which will not digest the cassette internally. The resulting PCR products are digested with the chosen restriction enzymes, and the fragments are ligated into a polylinker region of a standard cloning vector. The resulting plasmids are screened for orientation, and then the cassettes are released from the vector by digestion with appropriate restriction enzymes, filled in as necessary, and cloned into the polylinker region upstream of the *lacZ* gene in plasmid pZap2 as described in (a). Transformed *E. coli* colonies are plated on LBXIA medium as in (a), blue colonies are selected, and plasmid isolates are verified by restriction digestion and electrophoresis.

(h) ***Construction Additional Aspartokinase Expression Constructs Designed to Integrate***

Plasmid constructs designed to integrate into the *C. glutamicum* chromosome, which contain the *C. glutamicum* aspartokinase (*ask*) coding region operably associated with *C. glutamicum* regulatory regions contained in SEQ ID NOs 1, 2, and 4-22, are constructed as follows. Plasmid pBGS131 is digested with *Sma*I, or other appropriate restriction enzyme. The regulatory cassettes prepared as in (g) are released by digestion with appropriate restriction enzymes and the cassettes are subjected to fill-in reactions to produce blunt ends as needed. The cassettes are then ligated into the digested pBGS131. Plasmid isolates are tested for the proper orientation of the regulatory regions by restriction digestion and electrophoresis, to produce plasmids 131SEQ ID NO:1, 131 SEQ ID NO:2, and 131SEQ ID NO:4-131 SEQ ID NO:22. The *C. glutamicum ask* coding region is released from pTrec99A*ask* (produced as in (f)) by digestion with *Sac*I and *Kpn*I (or other appropriate restriction enzymes), or is PCR amplified from a *C. glutamicum* chromosome, and is cloned into the corresponding sites of plasmids 131SEQ ID NO:1, 131 SEQ ID NO:2, and 131SEQ ID NO:4-131 SEQ ID NO:22, to produce the plasmids 131SEQ ID NO:1-*ask*, 131 SEQ ID NO:2-*ask*, and 131SEQ ID NO:4-*ask* through 131 SEQ ID NO:22-*ask*.

(i) ***Construction of Additional Aspartokinase Constructs Designed to Replicate as a Plasmid***

Plasmid constructs designed to replicate as a plasmid in *C. glutamicum*, which contain the *C. glutamicum* aspartokinase (*ask*) coding region operably associated with *C. glutamicum* regulatory regions contained in SEQ ID NO1,2, and 4-22, are constructed as follows. Cassettes containing the *ask* coding region operably associated with the regulatory regions contained in SEQ ID NO:1, SEQ ID NO:2, and SEQ ID NO:4-SEQ ID NO:22 are released from plasmids 131SEQ ID NO:1-*ask*, 131 SEQ ID NO:2-*ask*, and 131SEQ ID NO:4-*ask* through 131

SEQ ID NO:22-*ask*, produced as in (h), by digestion with appropriate restriction enzymes. The resulting fragments are cloned into corresponding sites of plasmid pD10, which is capable of replicating in *C. glutamicum*, to produce plasmids PD10/SEQ ID NO:1-*ask*, PD10/SEQ ID NO:2-*ask*, and PD10/SEQ ID NO:4-*ask* through PD10/SEQ ID NO:22-*ask*.

Example 3

Use of LacZ as a Reporter to Demonstrate Transcriptional Regulation by araBAD (SEQ ID NO:23) and the lacI^q/trc Transcriptional Complex (SEQ ID NO:30) in Corynebacterium (41-004002)

The ability of the heterologous transcriptional regulatory regions lacI^q/trc and araBAD (with the araC repressor) to control gene expression in *C. glutamicum* was tested by measuring expression of β -galactosidase. Plasmids pTrcLacZ and pAraBADLacZ, produced as in Examples 2(a) and 2(b), respectively, were transferred into *C. glutamicum* according to the methods described in Example 1. The lacZ reporter gene, regulated by either of the two regulatory regions, was expressed from a plasmid replicating in *C. glutamicum*, grown in the presence of an inducer, IPTG in the case of the lacI^q/trc regulatory region, or arabinose in the case of the araC/araBAD regulatory region. The level of β -galactosidase production was measured, and the results are shown in Table 3.

Table 3

regulator/promoter	Inducer	nmol/min/mg	
		Mean	S.D.
none	none	0.40	0.14
LacI ^q /trc	none	3.84	0.57
LacI ^q /trc	+IPTG (30mg/L)	77.23	0.32
AraC/AraBAD	none	0.47	0.49
AraC/AraBAD	+arabinose (1 g/L)	5.51	0.46

The trc regulated construct shows uninduced activity. The araBAD regulated construct showed no uninduced activity and responded to arabinose. The

araBAD promoter resulted in a β -galactosidase expression level that was 14-fold lower than that observed with the *trc* promoter in this organism.

Example 4

Use of β -Galactosidase Activity to Demonstrate the Response of the C. glutamicum aceB Promoter to Acetate, Glucose and Ethanol (98-0040002)

The ability of the *C. glutamicum aceB* regulatory region (SEQ ID NO:3) to control gene expression of a heterologous reporter gene in *C. glutamicum* was tested by measuring expression of β -galactosidase from plasmid p*AceB10Zap2*, produced as in Example 2(c). In this construct, the *lacZ* reporter gene, regulated by *aceB*, is expressed from a plasmid replicating in *C. glutamicum*. Response of the regulatory region to various inducers was tested. The results are shown in Table 4.

Table 4

Rich Medium	nmol/min/mg
Minus sucrose	10
+ 2% glucose	3
+ 0.2% ammonium acetate	26
+ 2% ammonium acetate	50
+ 2% glucose + 0.2% ammonium acetate	7
+ 2% glucose + 2% ammonium acetate	15
+ 2% ethanol	28

Glucose repressed reporter activity. Ethanol induced activity but not as strongly as acetate. The uninduced control also showed activity implying that *aceB* is self-regulated in this medium.

Example 5

Use of the trc Promoter to Regulate Aspartokinase Activity when Integrated into the Chromosome of C. glutamicum (90-004002)

The ability of the *lacI^q/trc* (SEQ ID NO:30) regulatory region to control expression of aspartokinase, an enzyme in the *C. glutamicum* L-lysine biosynthesis pathway (encoded by the *ask* gene), was tested as follows. The 131-2 and 131-5 *C. glutamicum* strains have the *lacI^q/trc* transcriptional regulatory region, in operable association with the *ask* gene, integrated into the *C. glutamicum* chromosome. These strains were prepared by introducing plasmid 131LacI^q/trc-ask into *C. glutamicum* by the methods described in Example 1. The level of aspartokinase activity was measured both with and without the addition of the inducer IPTG. The results are shown in Table 5.

Table 5

Strain	Regulator/promoter-gene	Inducer	nmol/min/mg
131-2	<i>lacI^q/trc-ask</i>	none	59
131-2	<i>lacI^q/trc-ask</i>	+IPTG (30 mg/L)	117
131-5	<i>lacI^q/trc-ask</i>	none	59
131-5	<i>lacI^q/trc-ask</i>	+IPTG (30 mg/L)	123

The level of aspartokinase activity is doubled upon induction when the *trc* regulon is integrated into the chromosome.

Example 6

Lysine Production by Integrated, Uninduced lacI^q/trc-ask Constructs (84-004002)

This example shows the influence on lysine production of the uninduced *lacI^q/trc-ask* construct when integrated into the chromosome of *C. glutamicum*. The *lacI^q/trc* regulatory region was used to control expression of aspartokinase (encoded by the *ask* gene), an enzyme in the *C. glutamicum* L-lysine biosynthesis

pathway. Bacterial cultures of 131-2 and 131-5, prepared as described in Example 5, were grown in shaker flasks. BF100 is *C. glutamicum* strain which is a high level lysine producer. The level of aspartokinase activity was measured without the addition of an inducer. The results are shown in Table 6.

Table 6

Strain	Induction	O.D.	Titre	Yield	S.P.
BF100	none	43	23	39	53
131-2	none	34	27	46	82
131-5	none	35	28	47	82
O.D. = optical density at 660nm					
Titre = grams Lysine/liter					
Yield = grams lysine made/grams dextrose consumed					
S.P. = grams lysine/O.D.					

The leakiness of the *trc* promoter in this organism may be responsible for the improved productivity in the absence of an inducer.

Example 7

Improvement of Lysine Production in Shake Flasks by aceB-Aspartokinase Constructs with No Exogenous Induction (84-0040002)

This example shows the influence on lysine production of an *aceB*-Aspartokinase construct when integrated into the *C. glutamicum* chromosome. The *C. glutamicum* *aceB* regulatory region (SEQ ID NO:3) was used to control expression of aspartokinase (encoded by the *ask* gene), an enzyme in the *C. glutamicum* L-lysine biosynthesis pathway. Bacterial cultures were grown in shake flasks. Strain 131-6 is a *C. glutamicum* strain which has the *aceB* transcriptional regulatory region in operable association with the *ask* gene, integrated into the bacterial chromosome. Strain 131-6 was prepared by introducing plasmid 131*aceB/ask*, produced as described in Example 2(d), into *C. glutamicum* by the methods described in Example 1. BF100 is *C. glutamicum*

strain which is a high level lysine producer. The level of aspartokinase activity was measured either with or without the addition of the inducer acetate. The results are shown in Table 7.

Table 7

Strain	Induction	O.D.	Titre	Yield
BF100	none	43	25	37
	acetate	42	24	36
131-6	none	50	27	41
	acetate	47	27	40

Titer = g lysine/L medium

Yield = g lysine/ g glucose consumed * 100%

Induction by ammonium acetate is not required to observe an improvement in lysine production by this construct.

Example 8

Improvement of Lysine Production in Fermentors by aceB-aspartokinase Constructs with No Exogenous Induction (Lys002 Lys006RLK)

This example shows the influence on lysine production of uninduced aceB-Aspartokinase constructs which are expressed from a plasmid (PD10aceBask) replicating in *C. glutamicum*, or which are integrated into the *C. glutamicum* chromosome as in strain 131-6, described in Example 7. The *C. glutamicum* aceB regulatory region (SEQ ID NO:3) was used to control expression of aspartokinase (encoded by the ask gene), an enzyme in the *C. glutamicum* L-lysine biosynthesis pathway. Bacterial cultures were grown in fermentors. The *C. glutamicum* 131-aceB-aspartokinase strain, 131-6, has the aceB transcriptional regulatory region in operable association with the ask gene, integrated into the *C. glutamicum* chromosome. The *C. glutamicum* PD10-aceB-aspartokinase strain has the aceB transcriptional regulatory region in operable association with the ask gene, expressed from a plasmid (PD10aceBask, produced as described in Example 2(e)) replicating in *C. glutamicum*. BF100 is *C.*

-70-

glutamicum strain which is a high level lysine producer. Aspartokinase activity was measured in each of these three strains. The results are shown in Table 8.

Table 8

Strain	Titer	Yield	Product
BF100	127	32.9	735
131- <i>aceB</i> -aspartokinase	129	39.8	833
BF100	138	40.7	769
PD10- <i>aceB</i> -aspartokinase	135	45.8	830

Titer = g lysine/L medium

Yield = g lysine/g glucose consumed * 100%

Product = g lysine made in that fermented

The *aceB*-aspartokinase constructs contributed to productivity whether they were maintained on a plasmid or integrated into the chromosome. In neither case did the constructs need to be induced by an exogenous compound.

Example 9

Influence of Medium on Beta Galactosidase Activity Using Various Use of LacZ as a Reporter to Demonstrate Transcriptional Regulation by the Regulatory Regions Contained in SEQ ID NOs 1,2, and 4-22 in Corynebacterium

The ability of certain transcriptional regulatory regions disclosed herein to control gene expression in *C. glutamicum* was tested by measuring expression of β -galactosidase. The promoters used include (SEQ ID NO: 27) Tac, (SEQ ID NO: 1) Pta, (SEQ ID NO: 2) AceA, (SEQ ID NO: 3) AceB, (SEQ ID NO: 4) Adh, (SEQ ID NO: 5) AldB, (SEQ ID NO: 6) PoxB, (SEQ ID NO: 7) Ldh, (SEQ ID NO: 9) MalZ, (SEQ ID NO: 10) BglX, (SEQ ID NO: 13) HisD, (SEQ ID NO: 14) PyrR, (SEQ ID NO: 15) PurD, (SEQ ID NO: 18) DnaK, (SEQ ID NO: 20) GrpE, (SEQ ID NO: 21) ClpB, (SEQ ID NO: 33) LeuA, (SEQ ID NO: 34) IlvA and (SEQ ID NO: 35) IlvB. The reporter gene was expressed from plasmids replicating in *C. glutamicum*, and was driven by the regulatory regions contained in SEQ ID NOs 1,2, and 4-22. The plasmids were prepared as described in

Example 2(g). Increased expression of beta-galactosidase under the transcriptional control of these transcriptional regulatory regions is shown in Table 9. The "Seed" Medium is experimentally determined to support high growth. The "Main" Medium is formulated to induce higher expression of products, such as lysine, for example. Exemplary medias were described *supra*.

For certain regulatory regions, β -galactosidase levels are measured without the addition of any inducer or in the presence of various inducers, *e.g.*, acetic acid, pyruvate, ethanol, a starch subunit, a sugar, *e.g.*, fructose, maltose, lactose or arabinose, a cellulose subunit, a fatty acid, or a triglyceride. In addition, certain regulatory regions are tested by growth with a suppressor, *e.g.*, a purine, a pyrimidine, and amino acid, or oxygen, followed by a shift to medium lacking the suppressor at which time β -galactosidase levels were measured. For certain of the transcriptional regulatory regions, β -galactosidase levels are measured at optimal growth temperatures, and at temperatures either greater than or less than the optimal growth temperature. The β -galactosidase levels were measured using certain transcriptional regulatory regions at varying temperatures as shown in Table 10.

Table 9

Influence of Medium on beta-galactosidase Reporter Activity

(units = nmol/min/mg protein)

SEQ ID NO	Promoter	<u>Medium</u>		
		Seed(16hr)	Main(48hr)	Change
	Promoterless	2	20	19
27	Tac	231	342	111
1	Pta	134	111	-23
2	AceA	0	63	63
3	AceB	0	25	25
4	Adh	0	65	65

-72-

5	5	AldB	9	260	251
	6	PoxB	150	276	126
	7	Ldh	205	291	86
	9	MalZ	62	234	172
	10	BglX	57	208	151
10	13	HisD	119	241	122
	14	PyrR	161	289	128
	15	PurD	9	56	47
	18	DnaK	61	223	162
	20	GrpE	40	149	109
	21	ClpB	104	233	129
	33	LeuA	174	267	193
	34	IlvA	120	229	109
	35	IlvB	22	166	144

15

Table 10

Effect of Temperature Change on beta-galactosidase Reporter Activity
(units = nmol/min/mg protein; mean (standard deviation))
temperature shift in seed medium for 4 hrs

SEQ ID NO	Promoter	30°C	36°C	40 °C
20	Promoterless	0(0.7)	0(0.2)	0(0.7)
	21	clpB	76(7.3)	123(2.5)
	18	dnaK	70(0.8)	119(3.1)
	20	grpE	47(1.6)	107(1.6)
				97(2.6)

All 3 promoters increased reporter activity at 36°C

-73-

All 3 promoters increased reporter activity at 40°C

The *grpE* promoter doubled the activity at both 36 °C and 40 °C

Example 10

Improvement of Lysine Production Through Transcriptional Regulation of Aspartokinase by Regulatory Regions Contained in SEQ ID NOs 1,2, and 4-22

This example shows the influence on lysine production of induced or uninduced constructs in which the transcriptional regulatory regions of SEQ ID NOs 1,2, and 4-22 are used to express aspartokinase, either from a plasmid replicating in *C. glutamicum*, or when integrated into the *C. glutamicum* chromosome. The various *C. glutamicum* regulatory regions are used to control expression of aspartokinase, an enzyme in the *C. glutamicum* L-lysine biosynthesis pathway. Bacterial cultures are grown either in shake flasks or fermentors. *C. glutamicum* strains are prepared in which the transcriptional regulatory regions of SEQ ID NOs 1,2, and 4-22, in operable association with the *ask* gene, are integrated into the *C. glutamicum* chromosome. These strains are prepared by introducing plasmids 131 SEQ ID NO:1-*ask*, 131 SEQ ID NO:2-*ask*, and 131 SEQ ID NO:4-*ask* through 131 SEQ ID NO:22-*ask*, prepared as described in Example 2(h), into *C. glutamicum* by methods described in Example 1. Additional strains, with self-replicating plasmids are prepared by introducing plasmids PD10/SEQ ID NO:1-*ask*, PD10/SEQ ID NO:2-*ask*, and PD10/SEQ ID NO:4-*ask* through PD10/SEQ ID NO:22-*ask*, produced as described in Example 2(i), into *C. glutamicum* by methods described in Example 1. BF100, a high level lysine producer, is used as a control. Aspartokinase activity is measured in each of the resulting strains. For certain regulatory regions, aspartokinase levels are measured without the addition of any inducer, or in the presence of various inducers, *e.g.*, acetic acid, pyruvate, ethanol, a starch subunit, a sugar, *e.g.*, fructose, maltose, lactose or arabinose, a cellulose subunit, a fatty acid, or a

-74-

triglyceride. In addition, certain regulatory regions are tested by growth either without, or with a suppressor, *e.g.*, a purine, a pyrimidine, and amino acid, or oxygen, followed by a shift to medium lacking the suppressor at which time aspartokinase levels are measured. Finally, for certain regulatory regions, aspartokinase levels are measured at optimal growth temperatures, and also at temperatures either greater than or less than the optimal growth temperature.

* * * * *

Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that same can be performed by modifying or changing the invention with a wide and equivalent range of conditions, formulations and other parameters thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

What Is Claimed Is:

1. An isolated polynucleotide comprising a nucleic acid at least 90% identical to a reference nucleotide sequence selected from the group consisting of:

- 5 (a) SEQ ID NO: 4,
(b) SEQ ID NO: 5,
(c) SEQ ID NO: 6,
(d) SEQ ID NO: 7,
(e) SEQ ID NO: 8,
10 (f) SEQ ID NO: 9,
(g) SEQ ID NO: 10,
(h) SEQ ID NO: 11,
(i) SEQ ID NO: 12,
(j) SEQ ID NO: 13,
15 (k) SEQ ID NO: 14,
(l) SEQ ID NO: 15,
(m) SEQ ID NO: 16,
(n) SEQ ID NO: 17,
(o) SEQ ID NO: 18,
20 (p) SEQ ID NO: 19,
(q) SEQ ID NO: 20,
(r) SEQ ID NO: 21, and
(s) SEQ ID NO: 22.

25 2. The polynucleotide of claim 1, wherein said polynucleotide regulates transcription of β -galactosidase in a bacterial host cell.

3. The polynucleotide of claim 1, wherein the sequence of said nucleic acid is identical to said reference nucleotide sequence.

4. The polynucleotide of claim 1, further comprising a second nucleic acid.

5. The polynucleotide of claim 4, wherein said second nucleic acid encodes a polypeptide.

5 6. The polynucleotide of claim 5, wherein said polypeptide is selected from the group consisting of:

(a) a polypeptide which is a component of an amino acid biosynthesis pathway;

10 (b) a polypeptide which is a component of a purine nucleotide biosynthesis pathway; and

(c) a heterologous polypeptide.

7. The polynucleotide of claim 6, wherein said polypeptide is a component of an amino acid biosynthesis pathway.

15 8. The polynucleotide of claim 7 wherein said amino acid biosynthesis pathway is a lysine biosynthesis pathway.

9. The polynucleotide of claim 7, wherein said polypeptide is selected from the group consisting of:

(a) aspartokinase,

(b) diaminopimelate dehydrogenase,

20 (c) diaminopimelate decarboxylase,

(d) dihydrodipicolinate synthetase,

(e) dihydrodipicolinate reductase,

(f) aspartate beta-semialdehyde dehydrogenase, and

(g) pyruvate carboxylase.

-77-

10. A method of producing a vector which comprises inserting the polynucleotide of claim 1 into a vector.

11. A vector comprising the polynucleotide of claim 1.

12. A vector comprising the polynucleotide of claim 4.

5 13. A vector comprising the polynucleotide of claim 6.

14. A host cell comprising the vector of claim 11.

15. The host cell of claim 14, wherein said host cell is a *Corynebacterium* species.

10

16. A host cell comprising the vector of claim 12.

17. A host cell comprising the vector of claim 13.

18. A method of producing a transformed *Corynebacterium* species host cell comprising:

15

(a) introducing into *Corynebacterium* species cells the vector of claim 17, and

(b) selecting said host cell.

19. A method of production of a biosynthetic product, comprising culturing the host cell of claim 18 in or on a culture medium, and recovering said product.

20

-78-

20. An isolated polynucleotide comprising a nucleic acid which hybridizes to a reference nucleic acid, or the complement thereof, wherein the sequence of said reference nucleic acid is selected from the group consisting of:

- (a) SEQ ID NO: 4,
- 5 (b) SEQ ID NO: 5,
- (c) SEQ ID NO: 6,
- (d) SEQ ID NO: 7,
- (e) SEQ ID NO: 8,
- (f) SEQ ID NO: 9,
- 10 (g) SEQ ID NO: 10,
- (h) SEQ ID NO: 11,
- (i) SEQ ID NO: 12,
- (j) SEQ ID NO: 13,
- (k) SEQ ID NO: 14,
- 15 (l) SEQ ID NO: 15,
- (m) SEQ ID NO: 16,
- (n) SEQ ID NO: 17,
- (o) SEQ ID NO: 18,
- (p) SEQ ID NO: 19,
- 20 (q) SEQ ID NO: 20,
- (r) SEQ ID NO: 21, and
- (s) SEQ ID NO: 22;

wherein said first nucleic acid is at least 30 nucleotides in length.

21. The polynucleotide of claim 20, wherein said polynucleotide
25 regulates transcription of β -galactosidase in a bacterial host cell.

22. An isolated polynucleotide comprising a nucleic acid, the sequence of which comprises 10 contiguous nucleotides of a reference sequence selected from the group consisting of:

-79-

- 5 (a) SEQ ID NO: 4,
(b) SEQ ID NO: 5,
(c) SEQ ID NO: 6,
(d) SEQ ID NO: 7,
(e) SEQ ID NO: 8,
(f) SEQ ID NO: 9,
(g) SEQ ID NO: 10,
(h) SEQ ID NO: 11,
10 (i) SEQ ID NO: 12,
(j) SEQ ID NO: 13,
(k) SEQ ID NO: 14,
(l) SEQ ID NO: 15,
(m) SEQ ID NO: 16,
(n) SEQ ID NO: 17,
15 (o) SEQ ID NO: 18,
(p) SEQ ID NO: 19,
(q) SEQ ID NO: 20,
(r) SEQ ID NO: 21, and
(s) SEQ ID NO: 22.

20 23. The polynucleotide of claim 22, wherein said polynucleotide regulates transcription of β -galactosidase in a bacterial host cell.

24. The polynucleotide of claim 22, wherein the sequence of said first nucleic acid comprises 20 contiguous nucleotides of any of said sequences.

25 25. The polynucleotide of claim 24, wherein the sequence of said first nucleic acid comprises 50 contiguous nucleotides of any of said sequences.

26. The polynucleotide of claim 25, wherein the sequence of said first nucleic acid comprises 150 contiguous nucleotides of any of said sequences.

27. The polynucleotide of claim 24, further comprising a second nucleic acid.

5 28. The polynucleotide of claim 27, wherein said second nucleic acid encodes a polypeptide.

29. The polynucleotide of claim 28, wherein said polypeptide is selected from the group consisting of:

10 (a) a polypeptide which is a component of an amino acid biosynthesis pathway;

(b) a polypeptide which is a component of a purine nucleotide biosynthesis pathway; and

(c) a heterologous polypeptide.

15 30. The polynucleotide of claim 29, wherein said polypeptide is a component of an amino acid biosynthesis pathway.

31. A method of producing a vector which comprises inserting the polynucleotide of claim 22 into a vector.

32. A vector comprising the polynucleotide of claim 22.

33. A vector comprising the polynucleotide of claim 27.

20 34. A host cell comprising the vector of claim 32.

-81-

35. The host cell of claim 34, wherein said host cell is a *Corynebacterium* species.

36. A host cell comprising the vector of claim 33.

5 37. A method of producing a transformed *Corynebacterium* species host cell comprising:

(a) introducing into *Corynebacterium* species cells the vector of claim 33, and

(b) selecting said host cell.

10 38. A method of production of a biosynthetic product, comprising culturing the host cell of claim 36 in or on a culture medium, and recovering said product.

15 39. An isolated *Corynebacterium* species chromosome, comprising:
a first nucleic acid integrated into said chromosome, the sequence of which is at least 90% identical to SEQ ID NO:1, and
a second nucleic acid integrated into said chromosome, wherein said second nucleic acid encodes a polypeptide which functions as a component of an amino acid biosynthesis pathway; and
wherein said first nucleic acid regulates transcription of said second nucleic acid.

20 40. The *Corynebacterium* species chromosome of claim wherein the sequence of said first nucleic acid is identical to SEQ ID NO:1.

41. A host cell comprising the *Corynebacterium* species chromosome of claim 39.

-82-

42. A method of producing the *Corynebacterium* species host cell of claim 41, comprising:

(a) transforming *Corynebacterium* species cells with a vector comprising a first nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO: 1, wherein said vector facilitates integration of said first nucleic acid into the chromosome of said *Corynebacterium* species cells, and

(b) selecting said host cell.

43. A vector comprising the first and second polynucleotide of claim 39.

44. The vector of claim 43, wherein said second nucleic acid encodes a polypeptide which functions as a component of an amino acid biosynthesis pathway; and

wherein said first nucleic acid regulates transcription of said second nucleic acid.

45. A method of production of an amino acid, comprising culturing the host cell of claim 41 in or on a culture medium, and recovering said amino acid.

46. An isolated polynucleotide comprising:

a first nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO: 2, and

a second nucleic acid operably associated with said first nucleic acid, wherein said second nucleic acid encodes polypeptide which functions as a component of a lysine biosynthesis pathway;

wherein said first nucleic acid regulates transcription of said second nucleic acid.

-83-

47. The isolated polynucleotide of claim 46, wherein the sequence of said first nucleic acid is identical to SEQ ID NO:2.

48. A method of producing a vector, comprising inserting the polynucleotide of claim 47 into a vector.

5 49. A vector comprising the polynucleotide of claim 46.

50. A host cell comprising the vector of claim 49.

51. A method of producing a transformed *Corynebacterium* species host cell comprising:

- 10 (a) introducing into *Corynebacterium* species cells the vector of claim 49, and
(b) selecting said host cell.

52. An isolated polynucleotide comprising:
a first nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO: 3; and

15 a second nucleic acid operably associated with said second nucleic acid, wherein said second nucleic acid encodes polypeptide which functions as a component of an amino acid biosynthesis pathway;

wherein said first nucleic acid regulates transcription of said second nucleic acid.

20 53. The isolated polynucleotide of claim 52 wherein the sequence of said first nucleic acid is identical to SEQ ID NO:3.

54. A method of producing a vector, comprising inserting the polynucleotide of claim 52 into a vector.

55. A vector comprising the polynucleotide of claim 52.

56. A host cell comprising the vector of claim 55.

57. A method of producing a transformed *Corynebacterium* species host cell comprising:

- 5 (a) introducing into *Corynebacterium* species cells the vector of claim 55, and
(b) selecting said host cell.

58. A method of production of an amino acid, comprising culturing the host cell of claim 56 in or on a culture medium, and recovering said amino acid.
10

59. An isolated *Corynebacterium* species chromosome, comprising:
a first nucleic acid integrated into said chromosome, the sequence of which is at least 90% identical to SEQ ID NO:23, and
a second nucleic acid integrated into said chromosome in operable
15 association with said first nucleic acid, wherein said second nucleic acid encodes polypeptide which functions as a component of an amino acid biosynthesis pathway;
wherein said first nucleic acid regulates transcription of said second nucleic acid.

60. The *Corynebacterium* species chromosome of claim 59, wherein the sequence of said first nucleic acid is identical to SEQ ID NO:23.
20

61. A *Corynebacterium* species host cell comprising the *Corynebacterium* species chromosome of claim 59.

62. A method of producing the *Corynebacterium* species host cell of claim 61, comprising:

(a) transforming *Corynebacterium* species cells with a vector comprising a first nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO: 23, wherein said vector facilitates integration of said first nucleic acid into the chromosome of said *Corynebacterium* species cells, and

(b) selecting said host cell.

63. A vector comprising the *Corynebacterium* species chromosome of claim 59.

64. A method of production of an amino acid, comprising culturing the host cell of claim 61 in or on a culture medium, and recovering said amino acid.

65. The method of claim 64, further comprising adding arabinose to said culture medium.

66. The method of claim 65, wherein said first polynucleotide facilitates enhanced production of said amino acid in the presence of arabinose.

67. An isolated *Corynebacterium* species chromosome, comprising:
a first nucleic acid integrated into said chromosome, the sequence of which is identical to a nucleotide sequence selected from the group consisting of

(a) SEQ ID NO:26, and

(b) SEQ ID NO:27,

a second nucleic acid integrated into said chromosome in operable association with said first nucleic acid, wherein said second nucleic acid encodes

polypeptide which functions as a component of an amino acid biosynthesis pathway; and

a third nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO:28;

5 wherein said first nucleic acid regulates transcription of said second nucleic acid, and wherein said third nucleic acid encodes a lac repressor polypeptide, wherein said third nucleic acid is in operable association with said first nucleic acid.

10 68. The *Corynebacterium* species chromosome of claim 67, wherein the sequence of said third nucleic acid is identical to SEQ ID NO:28.

69. The *Corynebacterium* species chromosome of claim 67, wherein said third nucleic acid is operably associated with a transcriptional regulatory region.

15 70. A *Corynebacterium* species host cell comprising the *Corynebacterium* species chromosome of claim 67.

71. A method of producing the *Corynebacterium* species host cell of claim 70, comprising:

(a) transforming *Corynebacterium* species cells with a vector comprising:

20 (i) a first nucleic acid, the sequence of which is selected from the group consisting of SEQ ID NO:26 and SEQ ID NO:27; and

 (ii) a third nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO:28, operably associated with a transcriptional regulatory region, wherein said third nucleic acid encodes a lac repressor polypeptide; and

25

-87-

wherein said vector facilitates integration of said first and third nucleic acids into the chromosome of said *Corynebacterium* species cells, and

(b) selecting said host cell.

5 72. A vector comprising the first, second and third nucleic acid of claim 67.

 73. A method of production of an amino acid, comprising culturing the host cell of claim 70 in or on a culture medium, and recovering said amino acid.

-1-

SEQUENCE LISTING

<110> Rayapati, P. John
 Crafton, Corey M.

<120> Nucleotide Sequences for Transcriptional Regulation in
 Corynebacterium glutamicum

<130> 1533.194PC02

<140>

<141>

<150> 60/248,219

<151> 2000-11-15

<160> 35

<170> PatentIn version 3.0

<210> 1

<211> 500

<212> DNA

<213> Corynebacterium glutamicum

<400> 1

caccgacaac ggcaacacgc aaagggcgag acatataaag ttcgattcct taaaggggtt	60
ctaaaaaatg tggagtatgt gagcggggtt ccactagtag attcgactcc tatcggggtg	120
cgactgctaa tgggtgccctg ctatcaaccc tccatgatac gtggttaagt cagactaata	180
aaggccagtc ggggaggatt gggggctttg ctgggggcag atttgtcacg ctgcgcgctt	240
tcatagaccc cattaatgtg ggggtgaagag ctgtaaagta ccgctaataa ctttgcaaag	300
ggtgcttcgc aacttgtaac cgctccgtat tgttttctac ggcaataagc atttgtgctg	360
ctcaaagcgt ggaattgaga tcggtttgaa aattacaaaa taaaactttg caaacggggc	420
tgtacgcaag gcggacgaac gctaaactat gtaagaaatc acaacttccc ctcagtagtg	480
ccaggaggca caagcctgaa	500

-2-

<210> 2

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 2

```

actcttttaa gaaaagcact ctgactacct ctggaatcta ggtgccactc ttctttcgat      60
ttcaaccctt atcgtgtttg gcgatgtgat cagactaagt gatcacggtc accagcaaaa      120
ggggtttgcg aactttacta agtcattacc cccgcctaac cccgactttt atctagggtca      180
caccttcaaa acctacggaa cgttgcggtg cctgcatttt ccattttcag agcatttgcc      240
cagtacatct gtactagcaa ctccccgcc cactttttct gogaagccag aactttgcaa      300
acttcacaac aggggtgacc acccccgcac aaaacttaaa aacccaaacc gattgacgca      360
ccaatgcccg atggagcaat gtgtgaacca cgccaccacg caaacgatg cacatcacgt      420
cgaaacagtg acagtgcatt agctcatact ttgtggttgg caccgcccac tgogaatcag      480
cacttaagga agtgactttg                                     500

```

<210> 3

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 3

```

ttgcgtggtg gogtggttca cacattgctc catcgggcat tgggtgcgtca atcggtttgg      60
gtttttaagt tttgtgcggg ggtggtcacc cctgttgtga agtttgcaaa gttctggctt      120
cgcagaaaaa gtgggcgggg gagttgctag tacagatgta ctgggcaaata gctctgaaat      180
gggaaaatgc aggcacogca acgttcogta ggttttgaag gtgtgacctt gataaaagtc      240
ggggtttagc gggggtaatg acttagtaaa gttcgcaaac cctttttgct ggtgacggtg      300
atcacttagt ctgatcacat cgccaaacac gataaggggt gaaatcgaaa gaagagtggc      360
acctagattc cagaggtagt cagagtgcct ttcttaaaag agttttcaca accgttaacg      420
gcgtagccaa acaagaagga ttgcattct tctgggttag gcacaggtca tctaaaccca      480
tgctttaaaa ggagccttca                                     500

```

<210> 4

-3-

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 4

```

attggaaaat ggagattaga agaatcctgg gaatgttggt gtgtgttgca tgtctgttga      60
gactatctag tagatgcggc tgcgctcctt aattgcatgc tggggtggtg gggaatgggt      120
ggttgggggc gtcgaaaagc atttttgggtg cttctaaggg aattgtgtga atottggaaa      180
gctaattgaa aaacattccc attagtgggt gatttgctgg agttttgtga atctatTTTT      240
cgaaatttca gcgtgcgggg gtggtttgtt tttttacaat tgccagttca ttcaagggtg      300
ttgaaatgtt cgggggtaat aactcaactt tctatTTTca ccttgttggg atttcgctag      360
ggtggacgat ggcagcaatt gaatgttgta aatcacacaa ttgcaaggat tgtaatttaa      420
ggcacatcta tgcggtgtg aaattacatg tgccagaaga gcaatttgcc aagtaatcca      480
agcgagaagg agtgagtttt

```

<210> 5

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 5

```

cgatcatcga actcggcgaa gaaaaccgaa acctcaaaga atccctgcgt aagggtcacag      60
ctgagaatga gcagctcaaa gatcaattac gcagcgggcg tccgcgtggc gagctggtgc      120
acgtgccccg ctccaccgcg gtggtcatgt gggaacgccg caaggggcgt tccaagtaaa      180
aacatgcttg tcgacgccgc tttotagcaa attaagcggg cacctccatt tatotTTTtg      240
aggTgcccgt ttcgtgcttt cgccaattag atacatgcat aaccacccga acaggggtaa      300
taactTTTga aaggctttcg gcgttgagct gcgagaattt tgagaaaagg gggTgaattt      360
aacaggggtt ctagcgcgga ttgattttog tgaatatggt ggctgctaag cgtgcgaatg      420
tgcgcgttat cacaatcgtt gaccaagtgt cacctgacgc acaggtagtg ctcaggtgga      480
ggtggcccaa aggagacca

```

<210> 6

<211> 500

-4-

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 6

```

ttttagacca cggcgctgtg tggggattta agacgtcgga aattgtaggg gactgtcagt      60
gtgagtcggg ttctttgagg cgcttagagg cgattctgtg aggtcacttt ttgtggggtc     120
ggggctctaaa ttggccaggt ttctgaggcg accagacagg cgtgacaaga ttgactaaaa     180
aacogaagtt ttggcacgtg tgtttggttt ctoggggtct aaaccggaca ggcgtgacaa     240
gatctggcga aatcgaggt ttttgtcacg cctgtctggt ttacctttt gggggcccgga     300
actgcctga actactcgga tcgaccaagc agtttggcct ccagcgctct gatcaagcac     360
ccaaccgct ctaaatacaca ccaaggcact cgtaaaacc gtggcagata gagaaagtgt     420
ggcagcaact cgaattgaag agcacaattg aagtcgcacc aagttaggca acacaatagc     480
cataacgttg aggagttcag

```

<210> 7

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 7

```

aaaacagcca ggtagcggc tgtaaccac caoggtttcg gcaacaatga cggcgagaga      60
gccaccaca ttgcgatttc cgctccgata aagccagcgc ccatatttgc agggaggatt     120
cgctgcgggt ttggcgacat tcggatcccc ggaaccagct ctgcaatgac ctgcgcgcgcg     180
agggaaagcga ggtgggtggc aggttttagt gcgggtttta gcgttgccag gcgagtgggtg     240
agcaaagacg ctagtctggg gagcgaaacc atattgagtc atcttggcag agcatgcaca     300
attctgcagg gcatagattg gttttgctcg atttacaatg tgattttttc aacaaaaata     360
acaattggtc tgaccacatt ttcgacata atcgggcata attaaagggt taacaaagga     420
atccgggcac aagctcttgc tgattttctg agctgctttg tgggttgcc ggtagggaa     480
atcaggaagt gggatcgaaa

```

<210> 8

<211> 500

<212> DNA

-5-

<213> *Corynebacterium glutamicum*

<400> 8
ccagatctcg ggatgcttcg tagatcagca tggcgccgag gtcgttggct gctgcacgga 60
aagctgcgtt gtcgctgogc tcgtcgcgca acaggggtag gcggctagca acgagtgggt 120
ggttgacgat ggtgatgtcc atagcttcac atgttaaato attgccgcc cagaagaagac 180
cgcgcgggcg aatttgggct tggaggggaa caaacggcca cttttccagt ccaacaaagt 240
atgaggatta atttgcccca ctccaaagag ctgcgccacg agctgtgttt gttgccacc 300
cctgctgtgc ccgogcttcc cactgattct ggcgcgcagt ttgatatcca ccaggcactg 360
tccgcctctc ttgccaccta tgcccgcaac ctacacctgc tgtccacac cgccgagaat 420
ttaggaaacc gcgctctgac cggcctcgct gaaatcgaag acaccgacga ccaactcgca 480
cacgcattgg agcgctgac 500

<210> 9

<211> 1583

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 9
gggcatcatg ggcattgttg gtggattcac catgcctgtc gcaatcgcta ggaccagga 60
taagaacctc gtgtggttcc ccgtggctct tgggtgcatcg atgttcctcg gttatgtggg 120
aacgtggctg tggccgtccc aaggctggta cctgtggta ttccttcttg gtttaggtgg 180
actctgcttc ccgatggcta tcgcctgat tccagcgct acgaaagatc cgagaattac 240
cgcaagcttg tctggatttg tgcagccggt gggttacatt ctgcagccc ttgggccatt 300
ggcagtggga gcgatctacc aggcgattgg ctccgtgtca gagatcctcg ttggtttggc 360
cttgggcaca atagtgttgt cgattgtggg attcagagca gcacgcaatg tgacggttga 420
tgatgaattg aggagatcaa agtagcctca actaagcgtc gcgataagaa cgaggggcaa 480
ggctgatgta ctctgtcaac catggataaa ccggtcgtga gggatgcagc tctgctgatt 540
tttcgcgctg tgctcggagt gatctttgtg gcacacgggt gggaaaagct gttcatctcc 600
ggagttacca agacaacagg acaattttca gcctggggag tgctcaacc caagctctcg 660
gtgtggatca catogatctc tgagctgctc ggtgggtgct tcctagtggg ttggttgcctc 720
accacctttg ttgctgggtgc attagcgtg ttgatcgccg ctgctattta ctttgtgcac 780
ttgagttcgg gctttttcac agttgataac ggcacogaat tccccttgct catcattgtt 840

-6-

```

tctttgctcg tgatcgttgt gtttggttct ggtagagcca gcgttgatgg ggtgctcacg      900
cgtggttgac tgtagtgcca ttcaagccgc gctgtccgcc aaattagatg gtgagccgac      960
aggcttggat gatgcagtaa ttgaagcgca cctcgctaata tgtgaagagt gcagaaatta    1020
ctacaaccgt gctgctgagt tgaatcggat gctcaatttt tgcgccgcgg aacctcgcac    1080
cctgaccccg cctgatctat cagagatcat tctggcagag gtggaaccag aatggcgag      1140
gcatgccaac gccaaaggttg tgggatccct gctatcgcca gtgttggttg tgatcctggg      1200
tgtggtttac ctgcctggg gtatcacaat gttgggggat tcggcgctga taagcgtcca      1260
agaagaccgg ctacctcgcc gcctgctcgc ggaggccgtt gcctaccgca ttgctctttc      1320
tgtggggctg ttatttgccg cgtggaagcc gcggattatc gcgggcatgc tcccgatttt      1380
tggaacggtt tggacattta gtgctggttt tgctgcgcgc gatctcgtgt ttggcgctgc      1440
cgattcacag acgggactgt ccattggtct gctgttgatt tctacgattg tgctgtcggt      1500
tgcgttggtg aatagttctg gaccgggtat tttgcggcgc acatggaact cattgaacgc      1560
cgcgcccgcc taaggtggga ggc                                              1583

```

<210> 10

<211> 500

<212> DNA

<213> Corynebacterium glutamicum

<400> 10

```

taatacggca tccggcctga aaggacagtg tggaacccaa aaatcatgcc ctgcagaat      60
cgtgtttaag ggggttaaaac gcctcgaccc acactctgac ccatccgtga aactagacct      120
cttaaaacga cttctggttg gtgaggtgtg aaaaccgcac tgcggggccc aattcacaaa      180
aatccatacg agtgtgcaca cccgattttc acatcgcttc gagacctccc tttttgacac      240
cttaattgtc taaccccgta taggtgagaa atgttgagaa agtgtctgtt tttgtggggg      300
gaatctgact acgatggtaa gaaacaggga aaggggttac cattatgtct caagagtagc      360
ctcaaatcgg ctccgcctc tctcgtgtca ttgaacaaga cggcctacaa ttccgcgac      420
tcgacggcga cggcgacttt gcaccttatg aagattggcg tctaacccca gcagagcgtg      480
ccgctgacct ggtgaaacga                                              500

```

<210> 11

<211> 500

<212> DNA

-7-

<213> *Corynebacterium glutamicum*

<400> 11

```

caccgcgtcg ccgatctctc ccttccagac tcagaagttg cttoggcaat tgcacttgag      60
tgga aaccag gcgtcggatt cactacttga gctgctgatt tgtaggtttt aagaccttga      120
agatatagtt aattctcgtt gcaaggaacc agattccaag caatgcgtat tcctccatag      180
ctcagttggc agagcattcg actgttaatc gaagggtcac tggttcgagc ccagttggag      240
gagcaaattg aaaccactg ttttttaaca gtgggttttt ttgcatgttt catacagtta      300
acgaacccca attttctaac ttccactctc ctagctatga tgaatactcg ttgcaaggaa      360
gtaattcctt tccaatactt attcctccat agctcagttg gcagagcatt cgactgttaa      420
tcgaagggtc actggttcga gccagttgg aggagcaata cacacagccc gccgtttttc      480
ttaaaacggc gggttttgtt                                     500

```

<210> 12

<211> 1083

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 12

```

ggtgtcagtg tatccgcgaa ccagagcctc gatggcagct gtgtcttctg gggaagcggg      60
ctggccggtt tcaactggcg cagcttcagc gaaggtaatt gggttctcag acagctcttg      120
tgacagtaca tcaagctgtg cttgctgctc ttcgttgatg gttgttgctt cggaagattc      180
agattcagat tcagattcgg acgaggtggg ttctgctgct tccgcgagg tcgttcgct      240
ggaagaggaa gaatttgtcg tcgtcgctgc gctggatgag gctgcttctg tatcagatga      300
ctcactgcta catgcggtta agagaagtgg ggtgaccatg agtgctgca aagcagcctt      360
ctttgaggaa aggcgaatag acaaagttct gtcctgata aatcatcgac atgctccgga      420
aaacttaaaa attcccgac ggttcacgca gattacccta gcaaagcaat ctagctgacg      480
accaatttta gtctgtcat tatgctggca attgtgcagc tatctaaaga atctattatt      540
ggggcagccg tttcgatcct gagcgagttc ggtttgtcgg atatgaccat gcgtcgcgtc      600
gcaaagcaat taaatgtcgc gccgggcgcg ctgtattggc attttaaaaa taagcaggag      660
cttatcgacg ccacctcacg ccactctctg gcgcctatct tggggcgcaa cgacgagcag      720
cgagcaagca tttccgcgca ggaaacgtgc gcagaaatgc gttcactgat gatgcaaacc      780
aaagacggtg cggaagtcat cagtgccgca ctgagtaatc agcaactgcg ccaagaattg      840

```

-8-

```

gaatctctca tttccgactc tttaaaagaa cctaattgagg tcggtgcttt tacgctgctg      900
cattttgtgg tgggtgcagt attaacagaa caaactcagc tgcagatgca cgagttcacg      960
gctggcgcgg aagatgacac acaagaaaac cctgccgatg cgaactttga ggagagattc     1020
aatcaaggat tagaaatcat tctggcgggt ctagacgcgc ttgggcatat aagatagcgt      1080
tct                                                                    1083

```

<210> 13

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

```

<400> 13
agtgacagct cgcgccgcat cgttgatggc aaacatgtcc aaattaaggc gcatgcgacc      60
aaggatggty agcatttgcc ctgattcgtg gctgaagtac atcgaaatgc ggtgatcttg     120
ccacggcacg atgatgcgat cttctgagct caaatagtgg tagcccaagg aatcaacagc     180
ctccgtcact cgattcaggc caactgggaa aggaatggag gtgttgggtg cagggacggt     240
tcgatcattc acagaggtga aatccattcc ttccagtatc tcaaagggtga aagcggggtt     300
aattcaggta aatctggggg ggtcatttta agttttaagt ctaattcaaa tgaactctga     360
tgtaccctaaa tcagaaactt gttacgtggg gaatacaata ggtaaatatg cgggcttaag     420
aacttgtgtt gaggccgctt ggattcgggc accgagctcg aagaatttcg attcaacctt     480
ttaagggaga acttttcgcc                                                                    500

```

<210> 14

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

```

<400> 14
cgtggcgtgc caccatttta agtcccggcg ggagactgaa gatggtggty ccgtcggcga      60
ggcgttctgc ctgcgtaatg gggttaagtg ggatgaacgg gggagtcaga cgtgcgacag     120
cgcccttgcg ggtatgccaa tcccagacca tttctcgggg aaaaggaata aaatggcttg     180
tggtcagact cacaggggct tctccaagtc agtggattta tgaggtccca gtgggtacac     240
accgggtgtc ctacaacgat caattgtcac agattcgact ggcattgctgt accatctgct     300

```


-9-

ttaagcattt	tggtgtttca	ctgttggttaa	cagtgtttca	cogtggagca	ctaccttaga	360
tcatagtcag	catcttgggg	tgaatgtgac	acggtacgct	atagtgtcag	acaacaacca	420
ggaaactggt	cgttgcagag	tttttgcaaa	attggacatc	ctttaacgga	ccgcacagag	480
aggcggggaa	ggaggtcacg					500

<210> 15

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 15	
gccatggcgt	tgcggaatc gtaatcgcc attttgtcgg tggggaagag gtggatgtga 60
gtgtggggaa	catcgaatcc tgcgatgatg taaccacatc gaggggcgtc gaatgctgtg 120
cggattgcat	ttccgatgag ctgggaggcc tcgtttactt cgctccagat gttctgagga 180
aggtcgggcc	agcgggcaac ttctgcaacg ggtacgacta ggggtgtggcc gtaggtgagg 240
ggttcgatgg	atagaaaagc caccgacatc tcggaacgat acacaaatcg gccggggagc 300
tgcgcattaa	taattttcgt gaatacagaa gccatatgca cagactacta cttggcgtgc 360
aaccaaattg	aggttgcata aaataatgca ggtgagccgg tcattttaag gcgcttttcg 420
acgccacttt	caaccatttc cgaaccgcca agaatactgg aatagcttgg atcaagtttt 480
gcaggataaa	ctgtgcaacc 500

<210> 16

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 16	
acgctgatcg	ttttgagctt atcgacgctc gcctgcgctc agctggtttc gattggtacg 60
aggtatccaa	ctgggcgaaa cccggcggag aatgcaagca caacatgggc tattgggtcg 120
acggcgactg	gtggggtgct ggcccgggcg cgcactcgca catcggcgac cgccgcttct 180
acaacatcaa	gcaccagcgg cgttactccg cgcagattgc ggccggcgag ctgcccatta 240
aggaaacaga	gcggctgacg gcggaagatc accacaccga gcgcgtcatg cttggtttgc 300
gcctgaaaca	aggcgtgccg ctgaaccttt tcgcaccgcg agcgcgcccg gtcacgacc 360

-10-

gtcatatcgc aggaggcctg ctgcacgtca atgcgctggg caacctggcg gtgaccgatg 420
 cgggacgttt gottgccgac ggcatcatcg ccgacatttt gcttagtgaa gaagactaaa 480
 tatttagtag ggttacagac 500

<210> 17

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 17
 tggcgactat gccacatagt cgactacctt gcatagttga ctatttgatt gagttgaact 60
 gtgtcagttg atgaagagaa aaatgaagag aaaaacagct gcacatgggt tcaaaagata 120
 aggaagtga gcatgagcat cgagccagga atccccacgc ttggaccgct tgaagaacaa 180
 gtcatgcaca ttctgtggga tcacggaaaa ttgacagtcc gtgaagtcac cgaattcctt 240
 ccaggtgatc ctgcgtacac aacgatcgca accgtcctgc gtcacttggg cagaaaaggc 300
 atggtcacca ttgtgaaaga tggtcggact gtcgcacaca gcgcgttgat gaacagggaa 360
 gaatacacccg ctggcgctcat ggatcaggtg ctgtcgacca gtogggatcg caccgcatca 420
 attctgcatt tcgtggatac gatcacggcg actgatcgcg agctgcttct ggagtatctg 480
 caacagcagg agggcaggaa 500

<210> 18

<211> 297

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 18
 agacttggcg tcaaagttga gtgggacctg ttgaaaacag ctogtatctt ccctgattgg 60
 gtgggtgaaa ttaggggtaa acccggattt tttctcaagt gaagcgcttt gacctgtgta 120
 aattaagaaa gttgagtcta gtgggaacaa ctttgtggca tttaccgttg ccataatatgt 180
 aagcttgagt caggcaggct caatgaggag ttttcttac cggcgaaagt cgggtggaagc 240
 aagtcaaagc tcaagccgtg gacaatacta aaatcaccta aaacaggagg caccatt 297

<210> 19

<211> 500

-11-

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 19

```

cgggagccgt tcttctccac gatggggcgc tcggtgaaat cactgaaacg cgcgggtttcc      60
agaagcatca ccacgatggc gtcactgacc tcaaaagatg cagtgtgctc atcactgaag      120
acggtgtttt ctttgaagcc aagacctgca taaaaaacgc ttagatgctg caaggtcaga      180
tactggtagg ttgatgaaaa tcatgtcgtg ctgaagtogt gccatgggag ttccacttcc      240
tcaaagtgcc ttttggctaa tgtgacccca accagtatgc tgccgggttt tggattaggt      300
ttggccatcg tgctagcata tttaggtctc ggcgagggtc aagtactttt agtgetcaac      360
cgttatcgac gcgatctaga cttctaaagt gcacttttgt gcgctgcctg cgaagactcg      420
accaagacat tcgagtcggt cgcggggcatt ttttattttc ggggccgagt gtccaccttc      480
atccatgagg agaaatcact                                     500

```

<210> 20

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 20

```

tggtattgtct ggcggacaaa gttcgcaccg ataaaccogg caccaccggt cacaagcaaa      60
gaagtcattg tcgtcatagt agttgcggct cctgcagttg cgcgtgttat cgctgacggg      120
catgtcaatt cagccgtgac ccagcttcgg tgcggtgtga aaacgtatgg ggaaaagcca      180
cggatagtgc atgcccactt ctccacatgg gctgtgctca gctgtcgtag gggccacggc      240
ttgcggactt totgaacacc cggacgaaag cctccattat tgatggcacc caccgggggg      300
tggtatcggg tgctaagttc acagtggctg gagagaataa gtcagcactt ctgattacat      360
gtacgaaatc agggcgggtca aatgccagtt cccacccggc ttcactaagg cgggacgacc      420
atagaaacgc agctttctgac attgaatgcg ccggtttgct atggactaca ccgcatatga      480
ggaaagggct tgaaacgcac                                     500

```

<210> 21

<211> 500

<212> DNA

-12-

<213> *Corynebacterium glutamicum*

<400> 21

```

cttaatgagt acggcgtggc ttatatcggg ttcgctccgt ttgaaatogg catgaccatc      60
atcgcgccga tcgcggtgct cgcaggcttt actatggggg tgggttgggc gtcggttgatt    120
gttgcgatcg tgatttttgg cctcgcggtg ggtctgaagt ggttgccgga gcgcggacat    180
gtccgcggcg agggtaagcc gcaataaagg ttggaagcgc cgggtctagg tccggcgctt    240
ctttcgtacg cttttcgacg cctccctcca cgtaatatta aagttaaggg ttttccctga    300
tgcttaagtg gtagtcagtg cttaaaacttg actgcggtcc actcaattta ttttcaaatt    360
ttttgaactt gagtggaaca tactcaactc tttgtgcgtt atagatatta gagagttaaa    420
taatggcgct tgacctgcag gaaattgaga tcaacactga ttgtgtaggt tggcgcccaa    480
caaagaaagg gcgttgaaag                                     500

```

<210> 22

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 22

```

tggccattcc tgtggtccat gctcgccctg ttctttctca ctggactggg caacgccggc      60
acattcaaac aaatgcccac gattttgccc aaacgccaaag cagggtggcg gatcggtctg    120
accggtgcc a ttggtgcctt cggccccttc attgtcgggtg tottgccttc cttcactcca    180
actgtcgcgt tcttctgggg ctgcggtggtg ttcttcatca tcgccaccgc tttgacctgg    240
atctactacg cccgcccga cgtccattc ccgggataaa ccgaaaggcc aatccatgac    300
tacaactact tcttctggga agtcttctga acagtcttct gaaaagatca accccctctt    360
caagctcggc agtttcctaa gaaaaggcac cgtcggttct gaaggccagc agattttcct    420
tcagggcgga cgccaagccg atgtgttttt atcgcaaccc gatgggcggt cgataaaagt    480
cgtgcgctcc cacacatggc                                     500

```

<210> 23

<211> 1285

<212> DNA

<213> *Escherichia coli*

-13-

```

<400> 23
tgcataatgt gcctgtcaaa tggacgaagc agggattctg caaacccctat gctactccgt      60
caagccgtca attgtctgat tcgttaccaa ttatgacaac ttgacggcta catcattcac      120
ttttttttca caaccggcac ggaactcgct cgggctggcc ccggtgcatt ttttaaatac      180
ccgcgagaaa tagagttgat cgtcaaaacc aacattgcga ccgacggtag cgataggcat      240
ccgggtgggtg ctcaaaagca gcttcgcctg gctgatacgt tggtcctcgc gccagcttaa      300
gacgctaata cctaactgct ggcggaag atgtgacaga ccgacggcg acaagcaaac      360
atgtgtgtag acgctggcga tatcaaaatt gctgtctgcc aggtgatcgc tgatgtactg      420
acaagcctcg cgtacccgat tatccatcgg tggatggagc gactcgttaa tcgcttccat      480
gcgcgcagat aacaattgct caagcagatt tatcgccagc agctccgaat agcgcccttc      540
cccttgcccg gcgttaatga tttgccccaa caggctcgtg aaatgcggct ggtgcgcttc      600
atccggggcga aagaaccccg tattggcaaa tattgacggc cagttaagcc attcatgcca      660
gtaggcgcgc ggacgaaagt aaaccactg gtgataccat tcgcgagcct ccggatgacg      720
accgtagtga tgaatctctc ctggcgggaa cagcaaaata tcacccggtc ggcaaacaaa      780
ttctcgtccc tgatttttca ccacccctg accgcgaatg gtgagattga gaatataacc      840
tttcattccc agcggtcggt cgataaaaaa atcgagataa ccgttggcct caatcgcgct      900
taaaccgcgc accagatggg cattaaacga gtatcccgcc agcaggggat cattttgcgc      960
ttcagccata cttttcatac tcccgccatt cagagaagaa accaattgtc catattgcat     1020
cagacattgc cgtcactgcg tcttttactg gctcttctcg ctaaccaaac cggtaacccc     1080
gcttattaaa agcattctgt acaaagcgg gaccaagcc atgacaaaaa cgcgtaacaa     1140
aagtgtctat aatcacggca gaaaagtcca cattgattat ttgcacggcg tcacactttg     1200
ctatgccata gcatttttat ccataagatt agcggatcct acctgacgct ttttatcgca     1260
actctctact gttttcccat acccg                                     1285

```

<210> 24

<211> 1285

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (266)..(1192)

-14-

<400>	24																			
cg	ggt	at	gga	gaa	acag	tag	agag	ttg	cga	taaa	aag	cgt	cag	gtag	gat	ccg	ta	tct	60	
tat	ggata	aaa	aat	gct	at	gg	cat	ag	caa	aag	tgt	gac	gcc	tg	caa	ata	aat	caat	gt	120
tttt	ctg	ccg	tg	att	at	aga	cac	tttt	gtt	acg	cgt	tttt	gt	cat	gg	cct	tg	gt	ccc	180
tt	gttac	aga	at	gct	tttt	aa	taag	cgg	ggg	t	acc	gg	ttt	g	ttag	c	gag	acc	ag	240
aa	agac	gc	cag	tg	ac	gg	ca	at	gt	ctg	at	g	caa	t	at	g	caa	t	tg	292
ctg	aat	ggc	ggg	agt	atg	aaa	agt	atg	gct	gaa	gcg	caa	aat	gat	ccc					340
Leu	Asn	Gly	Gly	Ser	Met	Lys	Ser	Met	Ala	Glu	Ala	Gln	Asn	Asp	Pro					
10					15					20					25					
ctg	ctg	cgc	gga	tac	tcg	ttt	aat	gcc	cat	ctg	gtg	gcg	gg	tta	acg					388
Leu	Leu	Pro	Gly	Tyr	Ser	Phe	Asn	Ala	His	Leu	Val	Ala	Gly	Leu	Thr					
				30					35					40						
ccg	att	gag	gcc	aac	gg	t	at	ctc	gat	ttt	ttt	atc	gac	cga	ccg	ctg				436
Pro	Ile	Glu	Ala	Asn	Gly	Tyr	Leu	Asp	Phe	Phe	Ile	Asp	Arg	Pro	Leu					
			45					50					55							
gga	atg	aaa	gg	t	at	att	ctc	aat	ctc	acc	att	cgc	gg	cag	ggg	gtg				484
Gly	Met	Lys	Gly	Tyr	Ile	Leu	Asn	Leu	Thr	Ile	Arg	Gly	Gln	Gly	Val					
		60					65					70								
gtg	aaa	aat	cag	gga	cga	gaa	ttt	gtt	tgc	cga	ccg	gg	gat	att	ttg					532
Val	Lys	Asn	Gln	Gly	Arg	Glu	Phe	Val	Cys	Arg	Pro	Gly	Asp	Ile	Leu					
	75					80					85									
ctg	ttc	ccg	cca	gga	gag	att	cat	cac	tac	gg	cgt	cat	ccg	gag	gct					580
Leu	Phe	Pro	Pro	Gly	Glu	Ile	His	His	Tyr	Gly	Arg	His	Pro	Glu	Ala					
90					95					100					105					
cgc	gaa	tgg	tat	cac	cag	tgg	gtt	tac	ttt	cgt	ccg	cgc	gcc	tac	tgg					628
Arg	Glu	Trp	Tyr	His	Gln	Trp	Val	Tyr	Phe	Arg	Pro	Arg	Ala	Tyr	Trp					
				110					115					120						
cat	gaa	tgg	ctt	aac	tgg	ccg	tca	ata	ttt	gcc	aat	acg	ggg	ttc	ttt					676
His	Glu	Trp	Leu	Asn	Trp	Pro	Ser	Ile	Phe	Ala	Asn	Thr	Gly	Phe	Phe					
			125				130					135								
cgc	ccg	gat	gaa	gcg	cac	cag	ccg	cat	ttc	agc	gac	ctg	ttt	ggg	caa					724
Arg	Pro	Asp	Glu	Ala	His	Gln	Pro	His	Phe	Ser	Asp	Leu	Phe	Gly	Gln					
		140					145					150								
atc	att	aac	gcc	ggg	caa	ggg	gaa	ggg	cgc	tat	tcg	gag	ctg	ctg</						

-15-

tac atc agc gat cac ctg gca gac agc aat ttt gat atc gcc agc gtc 916
 Tyr Ile Ser Asp His Leu Ala Asp Ser Asn Phe Asp Ile Ala Ser Val
 205 210 215

gca cag cat gtt tgc ttg tcg ccg tcg cgt ctg tca cat ctt ttc cgc 964
 Ala Gln His Val Cys Leu Ser Pro Ser Arg Leu Ser His Leu Phe Arg
 220 225 230

cag cag tta ggg att agc gtc tta agc tgg cgc gag gac caa cgt atc 1012
 Gln Gln Leu Gly Ile Ser Val Leu Ser Trp Arg Glu Asp Gln Arg Ile
 235 240 245

agc cag gcg aag ctg ctt ttg agc acc acc cgg atg cct atc gcc acc 1060
 Ser Gln Ala Lys Leu Leu Leu Ser Thr Thr Arg Met Pro Ile Ala Thr
 250 255 260 265

gtc ggt cgc aat gtt ggt ttt gac gat caa ctc tat ttc tcg cgg gta 1108
 Val Gly Arg Asn Val Gly Phe Asp Asp Gln Leu Tyr Phe Ser Arg Val
 270 275 280

ttt aaa aaa tgc acc ggg gcc agc ccg agc gag ttc cgt gcc ggt tgt 1156
 Phe Lys Lys Cys Thr Gly Ala Ser Pro Ser Glu Phe Arg Ala Gly Cys
 285 290 295

gaa gaa aaa gtg aat gat gta gcc gtc aag ttg tca taattggttaa 1202
 Glu Glu Lys Val Asn Asp Val Ala Val Lys Leu Ser
 300 305

cgaatcagac aattgacggc ttgacggagt agcatagggt ttgcagaatc cctgcttcgt 1262
 ccatttgaca ggcacattat gca 1285

<210> 25

<211> 309

<212> PRT

<213> Escherichia coli

<400> 25

Met Gln Tyr Gly Gln Leu Val Ser Ser Leu Asn Gly Gly Ser Met Lys
1 5 10 15

Ser Met Ala Glu Ala Gln Asn Asp Pro Leu Leu Pro Gly Tyr Ser Phe
20 25 30

Asn Ala His Leu Val Ala Gly Leu Thr Pro Ile Glu Ala Asn Gly Tyr
35 40 45

Leu Asp Phe Phe Ile Asp Arg Pro Leu Gly Met Lys Gly Tyr Ile Leu
50 55 60

Asn Leu Thr Ile Arg Gly Gln Gly Val Val Lys Asn Gln Gly Arg Glu

-16-

65		70		75		80
Phe Val Cys Arg	Pro Gly Asp Ile Leu Leu Phe Pro Pro Gly Glu Ile	85		90		95
His His Tyr Gly Arg His Pro Glu Ala Arg Glu Trp Tyr His Gln Trp		100		105		110
Val Tyr Phe Arg Pro Arg Ala Tyr Trp His Glu Trp Leu Asn Trp Pro		115		120		125
Ser Ile Phe Ala Asn Thr Gly Phe Phe Arg Pro Asp Glu Ala His Gln		130		135		140
Pro His Phe Ser Asp Leu Phe Gly Gln Ile Ile Asn Ala Gly Gln Gly		145		150		155
Glu Gly Arg Tyr Ser Glu Leu Leu Ala Ile Asn Leu Leu Glu Gln Leu		165		170		175
Leu Leu Arg Arg Met Glu Ala Ile Asn Glu Ser Leu His Pro Pro Met		180		185		190
Asp Asn Arg Val Arg Glu Ala Cys Gln Tyr Ile Ser Asp His Leu Ala		195		200		205
Asp Ser Asn Phe Asp Ile Ala Ser Val Ala Gln His Val Cys Leu Ser		210		215		220
Pro Ser Arg Leu Ser His Leu Phe Arg Gln Gln Leu Gly Ile Ser Val		225		230		235
Leu Ser Trp Arg Glu Asp Gln Arg Ile Ser Gln Ala Lys Leu Leu Leu		245		250		255
Ser Thr Thr Arg Met Pro Ile Ala Thr Val Gly Arg Asn Val Gly Phe		260		265		270
Asp Asp Gln Leu Tyr Phe Ser Arg Val Phe Lys Lys Cys Thr Gly Ala		275		280		285
Ser Pro Ser Glu Phe Arg Ala Gly Cys Glu Glu Lys Val Asn Asp Val		290		295		300
Ala Val Lys Leu Ser		305				

-17-

<210> 26

<211> 56

<212> DNA

<213> Synthetic

<220>

<223> *lacI^q-trc* promoter

<400> 26

ttgacaatta atcatccggc tcgtataatg tgtggaattg tgagcggata acaatt 56

<210> 27

<211> 55

<212> DNA

<213> Synthetic

<220>

<223> *lacI^q-tac* promoter

<400> 27

ttgacaatta atcatcggct cgtataatgt gtggaattgt gagcggataa caatt 55

<210> 28

<211> 1092

<212> DNA

<213> *Escherichia coli*

<220>

<221> CDS

<222> (1)..(1089)

<400> 28

gtg	gtg	aat	gtg	aaa	cca	gta	acg	tta	tac	gat	gtc	gca	gag	tat	gcc	48
Val	Val	Asn	Val	Lys	Pro	Val	Thr	Leu	Tyr	Asp	Val	Ala	Glu	Tyr	Ala	
1				5					10					15		

ggc	gtc	tct	tat	cag	acc	gtt	tcc	cgc	gtg	gtg	aac	cag	gcc	agc	cac	96
Gly	Val	Ser	Tyr	Gln	Thr	Val	Ser	Arg	Val	Val	Asn	Gln	Ala	Ser	His	

-18-

20	25	30	
gtt tot gcg aaa acg cgg gaa aaa gtg gaa gcg gcg atg gcg gag ctg Val Ser Ala Lys Thr Arg Glu Lys Val Glu Ala Ala Met Ala Glu Leu 35 40 45			144
aat tac att ccc aac cgc gtg gca caa caa ctg gcg ggc aaa cag tcg Asn Tyr Ile Pro Asn Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser 50 55 60			192
ttg ctg att ggc gtt gcc acc tcc agt ctg gcc ctg cac gcg ccg tcg Leu Leu Ile Gly Val Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser 65 70 75 80			240
caa att gtc gcg gcg att aaa tct cgc gcc gat caa ctg ggt gcc agc Gln Ile Val Ala Ala Ile Lys Ser Arg Ala Asp Gln Leu Gly Ala Ser 85 90 95			288
gtg gtg gtg tct atg gta gaa cga agc ggc gtc gaa gcc tgt aaa gcg Val Val Val Ser Met Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala 100 105 110			336
gcg gtg cac aat ctt ctc gcg caa cgc gtc agt ggg ctg atc att aac Ala Val His Asn Leu Leu Ala Gln Arg Val Ser Gly Leu Ile Ile Asn 115 120 125			384
tat ccg ctg gat gac cag gat gcc att gct gtg gaa gct gcc tgc act Tyr Pro Leu Asp Asp Gln Asp Ala Ile Ala Val Glu Ala Ala Cys Thr 130 135 140			432
aat gtt ccg gcg tta ttt ctt gat gtc tct gac cag aca ccc atc aac Asn Val Pro Ala Leu Phe Leu Asp Val Ser Asp Gln Thr Pro Ile Asn 145 150 155 160			480
agt att att ttc tcc cat gaa gac ggt acg cga ctg ggc gtg gag cat Ser Ile Ile Phe Ser His Glu Asp Gly Thr Arg Leu Gly Val Glu His 165 170 175			528
ctg gtc gca ttg ggt cac cag caa atc gcg ctg tta gcg ggc cca tta Leu Val Ala Leu Gly His Gln Gln Ile Ala Leu Leu Ala Gly Pro Leu 180 185 190			576
agt tct gtc tct gcg cgt ctg cgt ctg gct ggc tgg cat aaa tat ctc Ser Ser Val Ser Ala Arg Leu Arg Leu Ala Gly Trp His Lys Tyr Leu 195 200 205			624
act cgc aat caa att cag ccg ata gcg gaa cgg gaa ggc gac tgg agt Thr Arg Asn Gln Ile Gln Pro Ile Ala Glu Arg Glu Gly Asp Trp Ser 210 215 220			672
gcc atg tcc ggt ttt caa caa acc atg caa atg ctg aat gag ggc atc Ala Met Ser Gly Phe Gln Gln Thr Met Gln Met Leu Asn Glu Gly Ile 225 230 235 240			720
gtt ccc act gcg atg ctg gtt gcc aac gat cag atg gcg ctg ggc gca Val Pro Thr Ala Met Leu Val Ala Asn Asp Gln Met Ala Leu Gly Ala 245 250 255			768
atg cgc gcc att acc gag tcc ggg ctg cgc gtt ggt gcg gat atc tcg Met Arg Ala Ile Thr Glu Ser Gly Leu Arg Val Gly Ala Asp Ile Ser 260 265 270			816

-19-

gta gtg gga tac gac gat acc gaa gac agc tca tgt tat atc ccg ccg 864
 Val Val Gly Tyr Asp Asp Thr Glu Asp Ser Ser Cys Tyr Ile Pro Pro
 275 280 285

tca acc acc atc aaa cag gat ttt cgc ctg ctg ggg caa acc agc gtg 912
 Ser Thr Thr Ile Lys Gln Asp Phe Arg Leu Leu Gly Gln Thr Ser Val
 290 295 300

gac cgc ttg ctg caa ctc tct cag ggc cag gcg gtg aag ggc aat cag 960
 Asp Arg Leu Leu Gln Leu Ser Gln Gly Gln Ala Val Lys Gly Asn Gln
 305 310 315 320

ctg ttg ccc gtc tca ctg gtg aaa aga aaa acc acc ctg gcg ccc aat 1008
 Leu Leu Pro Val Ser Leu Val Lys Arg Lys Thr Thr Leu Ala Pro Asn
 325 330 335

acg caa acc gcc tct ccc cgc gcg ttg gcc gat tca tta atg cag ctg 1056
 Thr Gln Thr Ala Ser Pro Arg Ala Leu Ala Asp Ser Leu Met Gln Leu
 340 345 350

gca cga cag gtt tcc cga ctg gaa agc ggg cag tga 1092
 Ala Arg Gln Val Ser Arg Leu Glu Ser Gly Gln

<210> 29

<211> 363

<212> PRT

<213> Escherichia coli

<400> 29

Val Val Asn Val Lys Pro Val Thr Leu Tyr Asp Val Ala Glu Tyr Ala
 1 5 10 15

Gly Val Ser Tyr Gln Thr Val Ser Arg Val Val Asn Gln Ala Ser His
 20 25 30

Val Ser Ala Lys Thr Arg Glu Lys Val Glu Ala Ala Met Ala Glu Leu
 35 40 45

Asn Tyr Ile Pro Asn Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser
 50 55 60

Leu Leu Ile Gly Val Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser
 65 70 75 80

Gln Ile Val Ala Ala Ile Lys Ser Arg Ala Asp Gln Leu Gly Ala Ser
 85 90 95

Val Val Val Ser Met Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala

-20-

100					105					110					
Ala	Val	His	Asn	Leu	Leu	Ala	Gln	Arg	Val	Ser	Gly	Leu	Ile	Ile	Asn
	115						120					125			
Tyr	Pro	Leu	Asp	Asp	Gln	Asp	Ala	Ile	Ala	Val	Glu	Ala	Ala	Cys	Thr
	130					135					140				
Asn	Val	Pro	Ala	Leu	Phe	Leu	Asp	Val	Ser	Asp	Gln	Thr	Pro	Ile	Asn
145					150					155					160
Ser	Ile	Ile	Phe	Ser	His	Glu	Asp	Gly	Thr	Arg	Leu	Gly	Val	Glu	His
				165					170					175	
Leu	Val	Ala	Leu	Gly	His	Gln	Gln	Ile	Ala	Leu	Leu	Ala	Gly	Pro	Leu
			180					185					190		
Ser	Ser	Val	Ser	Ala	Arg	Leu	Arg	Leu	Ala	Gly	Trp	His	Lys	Tyr	Leu
		195					200					205			
Thr	Arg	Asn	Gln	Ile	Gln	Pro	Ile	Ala	Glu	Arg	Glu	Gly	Asp	Trp	Ser
	210					215					220				
Ala	Met	Ser	Gly	Phe	Gln	Gln	Thr	Met	Gln	Met	Leu	Asn	Glu	Gly	Ile
225					230					235					240
Val	Pro	Thr	Ala	Met	Leu	Val	Ala	Asn	Asp	Gln	Met	Ala	Leu	Gly	Ala
				245					250					255	
Met	Arg	Ala	Ile	Thr	Glu	Ser	Gly	Leu	Arg	Val	Gly	Ala	Asp	Ile	Ser
			260					265					270		
Val	Val	Gly	Tyr	Asp	Asp	Thr	Glu	Asp	Ser	Ser	Cys	Tyr	Ile	Pro	Pro
		275					280					285			
Ser	Thr	Thr	Ile	Lys	Gln	Asp	Phe	Arg	Leu	Leu	Gly	Gln	Thr	Ser	Val
	290					295					300				
Asp	Arg	Leu	Leu	Gln	Leu	Ser	Gln	Gly	Gln	Ala	Val	Lys	Gly	Asn	Gln
305				310						315					320
Leu	Leu	Pro	Val	Ser	Leu	Val	Lys	Arg	Lys	Thr	Thr	Leu	Ala	Pro	Asn
				325					330					335	
Thr	Gln	Thr	Ala	Ser	Pro	Arg	Ala	Leu	Ala	Asp	Ser	Leu	Met	Gln	Leu
			340					345					350		

-21-

Ala Arg Gln Val Ser Arg Leu Glu Ser Gly Gln
 355 360

<210> 30

<211> 1511

<212> DNA

<213> Escherichia coli

<400> 30

```

catgcattta cgttgacacc atcgaatggt gcaaaacott tcgcggtatg gcatgatagc      60
gcccggaaga gagtcaattc aggggtggtga atgtgaaacc agtaacgtta tacgatgtcg    120
cagagtatgc cgggtgtctct tatcagaccg tttcccgcggt ggtgaaccag gccagccacg    180
tttctgcgaa aacgcgggaa aaagtggaag cggcgatggc ggagctgaat tacattccca    240
accgcgtggc acaacaactg gcggggcaaac agtcgttget gattggcggt gccacctcca    300
gtctggccct gcacgcgccg tcgcaaattg tcgcggcgat taaatctcgc gccgatcaac     360
tgggtgccag cgtggtggtg tcgatggtag aacgaagcgg cgtcgaagcc tgtaaagcgg     420
cgggtgcacaa tcttctcgcg caacgcgtca gtgggctgat cattaactat ccgctggatg     480
accaggatgc cattgctgtg gaagctgcct gcactaatgt tccggcgta tttcttgatg     540
tctctgacca gacaccatc aacagtatta tttctccca tgaagacggt acgcgactgg     600
gcgtggagca tctggtcgca ttgggtcacc agcaaatcgc gctgtagcg ggcccattaa     660
gttctgtctc ggcgcgctct cgtctggctg gctggcataa atatctcact cgcaatcaaa     720
ttcagccgat agcggaacgg gaaggcgact ggagtgocat gtccggtttt caacaaacca     780
tgcaaatgct gaatgagggc atcgttccca ctgcgatgct ggttgccaac gatcagatgg     840
cgctgggcgc aatgcgcgcc attaccgagt ccgggctgcg cgttggtgcg gatattctcg     900
tagtgggata cgacgatacc gaagacagct catgttatat cccgccgtca accaccatca     960
aacaggatth tcgcctgctg gggcaaacca gcgtggaccg cttgctgcaa ctctctcagg    1020
gccaggcggt gaagggcaat cagctgttgc ccgtctcact ggtgaaaaga aaaaccaccc    1080
tggcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg    1140
cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag    1200
cgcgaattga tctggtttga cagcttatca tcgactgcac ggtgcaccaa tgcttctggc    1260
gtcaggcagc catcggaagc tgtggtatgg ctgtgcaggt cgtaaatac tcataaatc     1320
gtgtcgctca aggcgcactc ccgttctgga taatgttttt tgcgccgaca tcataacggt    1380
tctggcaaat attctgaaat gagctgttga caattaatca tccggctcgt ataatgtgtg    1440

```

-22-

gaattgtgag cggataacaa tttcacacag gaaacagacc atggaattcg agctcggtac 1500
 ccggggatcc t 1511

<210> 31

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotide forward primer

<400> 31
 gtgcggatcc tggctttcca acgttt 26

<210> 32

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Oligonucleotide reverse primer

<400> 32
 catgggctcg agatgacctg tgccta 26

<210> 33

<211> 544

<212> DNA

<213> Corynebacterium glutamicum

<400> 33
 ctgggaaggg taccgcctaa tctcaccgc attacttggg acaatcgtgc agcagtattc 60
 ctttaacgct ggcgaactac aaaaatcgct acccgccatg accattgccg aaccaattgt 120
 cgccttcagt ttgggctact tggttctggg cgaaaaattc caagtcgtgg actgggaatg 180
 gatcgccatg ggcacgcac tottgggtgat gattgtttcc accattgcac tgtctcgtac 240
 aagcacaatg ccggccgat cgaaaaggta aaactccaaa gttccccccc cgagacatga 300

-23-

cagcactgga actgggcgtc gaaaagcttt tttaaaagaa aactcccccc gagttgctac	360
ccacaccaca aagttgttgt atgcttcacc acatgacttc gcgtgcgaat ctacttcttc	420
ttcgccgcgg cggttcccag aggtcttaac acgaccggca tcccgtcgcg gagtttggtg	480
ttgccggtcg tggaccacc caaaactttt taagaaggtt ctagacaatg tcttctaacg	540
atgc	544

<210> 34

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 34

gaaccaatca ctggtacctc ggacaaccga tcgatcagtg cttgctgctc ttcgctgatt	60
tcacgtcat catcgctggg gtctttgccg gaggagctca agctcggggc cgcaggaatg	120
tcgccactgc tgagcattga gctgccttca gagctgcctg gccaggtttc gtttccatcg	180
actggttttc catcatcatc aaggatctgt gatgagggtga tgttgtctga gagctgtgtc	240
agtgcgtcag aggactgagc ctgggcaact ggagtgaaca cggacaatgc cacagcgctt	300
gctgtaacaa gggtaaaagt acttcgacgc aaagacaaaa cttttctcct ggcaataaat	360
atgcggatth actatggaaa caagatagaa gattggatag cgaaagctat cctcaactcg	420
tggaaagtgt aatgccacaa ccacagtatt ggctagaaaa caatctatag cattgttcta	480
caaagagctc gttggaaata	500

<210> 35

<211> 500

<212> DNA

<213> *Corynebacterium glutamicum*

<400> 35

ctacatcaaa accggtaccg acaatccaat tccacaatga atagagcaaa tattgaatgg	60
gtacgcctaa aatcatgagc caagattagc gctgaaaagt agcgggagcc tgccatgaact	120
ttgtgagaat cctgattcct taaccgaagt gggggagttt tgggggtggg aattttcgtg	180
cgttgtggaa ttggaaactc gatgtgtgta gcatgacaca ccatgaccat tattcgactt	240
gtagtagtaa ccgcgcggcg cctgcgctaa cggccttcca agtcgtctcg tcaagcgccc	300

-24-

tcgacaacac tcaccacagt gttggaacga gggctttctt gttgggttatg acccaagtag	360
ccaactttgc aacagacatc tgcgcactg cgtgcacacg catccgcgtc ggaacaattt	420
taaatgaggg ctttgtcttt aggctgagtt gaaatcggct tggcttggac gggtcctgtg	480
aaaatcctta tttagtaaag	500

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
23 May 2002 (23.05.2002)

PCT

(10) International Publication Number
WO 02/040679 A3

(51) International Patent Classification⁷: **C12N 15/31**,
15/62, 15/77, 1/21, C12P 13/04

(21) International Application Number: PCT/US01/43096

(22) International Filing Date:
15 November 2001 (15.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/248,219 15 November 2000 (15.11.2000) US

(71) Applicant: **ARCHER-DANIELS-MIDLAND COM-
PANY** [US/US]; 4666 Fairies Parkway, Box 1470,
Decatur, IL 62526 (US).

(72) Inventors: **RAYAPATI, P. John**; 601 North Union Street,
Monticello, IL 61856 (US). **CRAFTON, Corey, M.**; 1371
West Pershing Road, Decatur, IL 62526 (US).

(74) Agents: **LUDWIG, Steven, R.** et al.; Sterne, Kessler,
Goldstein & Fox P.L.L.C., 1100 New York Avenue, N.W.,
Suite 600, Washington, DC 20005-3934 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU,
ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
8 May 2003

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 02/040679 A3

(54) Title: CORYNEBACTERIUM GLUTAMICUM PROMOTERS

(57) Abstract: The invention relates to isolated polynucleotides from *Corynebacterium glutamicum* which are useful in the regulation of gene expression. In particular, the invention relates to isolated polynucleotides comprising *C. glutamicum* promoters which may be used to regulate, i.e., either increase or decrease, gene expression. In certain embodiments, isolated promoter sequences of the present invention regulate gene expression through the use of exogenous or endogenous induction. The invention further provides recombinant vectors and recombinant cells comprising isolated polynucleotides of the present invention, preferably in operable association with heterologous genes. Also provided are methods of regulating bacterial gene expression comprising growth of a recombinant cell of the present invention. In particular, the present invention provides methods to regulate genes involved in amino acid production comprising growth of a recombinant cell of the present invention. In certain embodiments, the present invention provides methods of regulating gene expression in bacteria, particularly *Corynebacterium* species, especially of the genus *Corynebacterium*, comprising fermentation growth of a recombinant cell of the present invention, where metabolite concentrations, temperature, or oxygen levels are manipulated to regulate gene expression.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/43096

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/31 C12N15/62 C12N15/77 C12N1/21 C12P13/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No. .
P,X	EP 1 108 790 A (KYOWA HAKKO KOGYO KK) 20 June 2001 (2001-06-20) SEQ ID NO:7068 ---	1,20
P,X	WO 01 00844 A (BASF AG) 4 January 2001 (2001-01-04) SEQ ID NO:209 ---	20
A	WO 96 15246 A (KERNFORSCHUNGSANLAGE JUELICH ;REINSCHIED DIETER (CA); EIKMANNS BER) 23 May 1996 (1996-05-23) page 5, line 10 -page 7, line 4 --- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

11 September 2002

Date of mailing of the international search report

09. 12. 2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Mata-Vicente, M

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 01/43096

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>PATEK MIROSLAV ET AL: "Promoters from <i>Corynebacterium glutamicum</i>: Cloning, molecular analysis and search for a consensus motif."</p> <p>MICROBIOLOGY (READING), vol. 142, no. 5, 1996, pages 1297-1309, XP008006983 ISSN: 1350-0872 page 1298, left-hand column, paragraph 3 figure 1 page 1307, right-hand column, paragraph 2</p> <p>---</p>	
A	<p>VASICOVA P ET AL: "Integrative and autonomously replicating vectors for analysis of promoters in <i>Corynebacterium glutamicum</i>."</p> <p>BIOTECHNOLOGY TECHNIQUES, vol. 12, no. 10, October 1998 (1998-10), pages 743-746, XP008006984 ISSN: 0951-208X page 745, left-hand column, paragraph 1</p> <p>---</p>	
A	<p>VASICOVA PAVLA ET AL: "Analysis of the <i>Corynebacterium glutamicum</i> dapA promoter."</p> <p>JOURNAL OF BACTERIOLOGY, vol. 181, no. 19, October 1999 (1999-10), pages 6188-6191, XP002211212 ISSN: 0021-9193 abstract</p> <p>---</p>	
A	<p>WO 88 09819 A (MASSACHUSETTS INST TECHNOLOGY) 15 December 1988 (1988-12-15) page 3, paragraph 3</p> <p>---</p>	
A	<p>REINSCHIED DIETER J ET AL: "Cloning, sequence analysis, expression and inactivation of the <i>Corynebacterium glutamicum</i> pta-ack operon encoding phosphotransacetylase and acetate kinase."</p> <p>MICROBIOLOGY (READING), vol. 145, no. 2, February 1999 (1999-02), pages 503-513, XP002211213 ISSN: 1350-0872 abstract</p> <p>-----</p>	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/43096

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-38 partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: Claims (1-38) - partially

Isolated *Corynebacterium glutamicum* promoter comprising SEQ ID NO:4, homologs and fragments; fusions, vectors containing them, host cells carrying the vectors and use of said cells for production of biosynthetic products.

Inventions 2-19: Claims (1-38) - partially

Idem as invention 1, but restricted to SEQ ID NOs:5-22 respectively.

Invention 20: Claims (39-45) - completely

Isolated *Corynebacterium* species chromosome comprising: a first nucleic acid integrated into said chromosome, the sequence of which is at least 90% identical to SEQ ID NO:1, and a second nucleic acid integrated into said chromosome, wherein said second nucleic acid encodes a polypeptide which functions as a component of an amino acid biosynthesis pathway; and wherein said first nucleic acid regulates transcription of said second nucleic acid. Host cell comprising said chromosome and use thereof to produce an amino acid. Vector comprising the first and second polynucleotides mentioned above.

Invention 21: Claims (46-51) - completely

Isolated polynucleotide comprising: a first nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO:2, and a second nucleic acid operably associated with said first nucleic acid, wherein said second nucleic acid encodes a polypeptide which functions as a component of a lysine biosynthesis pathway; wherein said first nucleic acid regulates transcription of said second nucleic acid. Vector comprising the polynucleotide and host cell transformed with it.

Invention 22: Claims (52-58) - completely

Isolated polynucleotide comprising: a first nucleic acid, the sequence of which is at least 90% identical to SEQ ID NO:3, and a second nucleic acid operably associated with said first nucleic acid, wherein said second nucleic acid encodes a polypeptide which functions as a component of an amino acid biosynthesis pathway; wherein said first nucleic acid regulates transcription of said second nucleic acid. Vector comprising the polynucleotide, host cell transformed

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

with it and use of said transformant to produce an amino acid.

Invention 23: Claims (59-66) - completely

Isolated *Corynebacterium* species chromosome comprising: a first nucleic acid integrated into said chromosome, the sequence of which is at least 90% identical to SEQ ID NO:23, and a second nucleic acid integrated into said chromosome in operable association with said first nucleic acid, wherein said second nucleic acid encodes a polypeptide which functions as a component of an amino acid biosynthesis pathway; and wherein said first nucleic acid regulates transcription of said second nucleic acid. Host cell comprising said chromosome and use thereof to produce an amino acid. Vector comprising the first and second polynucleotides mentioned above.

Invention 24: Claims (67-73) - completely

Isolated *Corynebacterium* species chromosome comprising: a first nucleic acid integrated into said chromosome, the sequence of which is identical to SEQ ID NO:26 or to SEQ ID NO:27, a second nucleic acid integrated into said chromosome in operable association with said first nucleic acid, wherein said second nucleic acid encodes a polypeptide which functions as a component of an amino acid biosynthesis pathway, and a third nucleic acid, operably associated to the first nucleic acid and the sequence of which is at least 90% identical to SEQ ID NO:28; wherein said first nucleic acid regulates transcription of said second nucleic acid, and wherein said third nucleic acid encodes a lac repressor polypeptide. Host cell comprising said chromosome and use thereof to produce an amino acid. Vector comprising the first, the second and the third polynucleotides mentioned above.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/43096

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 1108790	A	20-06-2001	EP 1108790 A2	20-06-2001
			JP 2002191370 A	09-07-2002

WO 0100844	A	04-01-2001	AU 5559000 A	31-01-2001
			CN 1370235 T	18-09-2002
			WO 0100844 A2	04-01-2001
			TR 200103706 T2	21-10-2002
			BR 0011805 A	14-05-2002
			EP 1257649 A2	20-11-2002
			WO 0100843 A2	04-01-2001
			TR 200103707 T2	23-09-2002
			AU 5836900 A	31-01-2001
			BR 0011811 A	18-06-2002
			WO 0100804 A2	04-01-2001
			SK 18882001 A3	10-09-2002
			TR 200103709 T2	21-08-2002
			AU 5421300 A	31-01-2001
			BR 0011806 A	14-05-2002
			CN 1371417 T	25-09-2002
			AU 5421600 A	31-01-2001
			BR 0011810 A	07-05-2002
			CN 1370236 T	18-09-2002
			EP 1255839 A2	13-11-2002
			WO 0100805 A2	04-01-2001
			TR 200103708 T2	21-08-2002
			AU 5420500 A	31-01-2001
			EP 1254232 A2	06-11-2002
			WO 0100842 A2	04-01-2001
			SK 18902001 A3	10-09-2002
			TR 200103711 T2	22-07-2002
			CN 1370234 T	18-09-2002
			AU 5701400 A	22-01-2001
			BR 0012038 A	02-07-2002
			CN 1371420 T	25-09-2002
			CZ 20014700 A3	12-06-2002
			EP 1246922 A2	09-10-2002
			WO 0102583 A2	11-01-2001
			SK 18892001 A3	10-09-2002
			TR 200103854 T2	21-06-2002
			BR 0011803 A	09-04-2002

WO 9615246	A	23-05-1996	DE 4440118 C1	09-11-1995
			CN 1174574 A	25-02-1998
			WO 9615246 A1	23-05-1996
			EP 0791062 A1	27-08-1997
			JP 10512742 T	08-12-1998
			US 5965391 A	12-10-1999
			ZA 9509598 A	28-05-1996

WO 8809819	A	15-12-1988	WO 8809819 A2	15-12-1988
			US 5641660 A	24-06-1997
